

日本国特許庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出願年月日

Date of Application: 2002年10月30日

出願番号

Application Number: 特願2002-316807

[ST.10/C]:

[JP2002-316807]

出願人

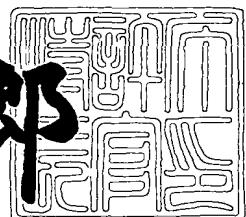
Applicant(s): ヤマハ株式会社

（2.

2003年 6月23日

特許庁長官
Commissioner,
Japan Patent Office

太田 信一郎



出証番号 出証特2003-3048492

【書類名】 特許願

【整理番号】 C30766

【提出日】 平成14年10月30日

【あて先】 特許庁長官 殿

【国際特許分類】 G11B 31/00

【発明の名称】 音声データと演奏データの同期再生を行うための装置および方法

【請求項の数】 7

【発明者】

【住所又は居所】 静岡県浜松市中沢町10番1号 ヤマハ株式会社内

【氏名】 古川 令

【特許出願人】

【識別番号】 000004075

【氏名又は名称】 ヤマハ株式会社

【代理人】

【識別番号】 100098084

【弁理士】

【氏名又は名称】 川▲崎▼ 研二

【手数料の表示】

【予納台帳番号】 038265

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【ブルーフの要否】 要

【書類名】 明細書

【発明の名称】 音声データと演奏データの同期再生を行うための装置および方法

【特許請求の範囲】

【請求項1】 演奏の制御を指示する制御データおよび該演奏の制御の実行タイミングを指示する演奏タイミングデータからなる演奏データを受信する演奏データ受信手段と、

樂曲の音声波形を示す音声データを受信する音声データ受信手段と、

前記音声データから、前記樂曲の一部の音声波形を示す第1部分データを取り出す第1部分データ取得手段と、

前記第1部分データが表す音声波形を抽象化した第1参照用データを生成する第1参照用データ生成手段と、

前記音声データから、前記樂曲の一部とは異なる音声波形を示す第2部分データを取り出す第2部分データ取得手段と、

前記第2部分データが表す音声波形を抽象化した第2参照用データを生成する第2参照用データ生成手段と、

前記第1参照用データ、前記第1部分データの再生タイミングを示す第1基準タイミングデータ、前記第2参照用データおよび前記第2部分データの再生タイミングを示す第2基準タイミングデータを、前記演奏データと共に記録する記録手段と

を備えることを特徴とする記録装置。

【請求項2】 演奏の制御を指示する制御データおよび該演奏の制御の実行タイミングを指示する演奏タイミングデータからなる演奏データを受信する演奏データ受信手段と、

樂曲の音声波形の一部を抽象化した第1参照用データおよび前記第1参照用データに対応づけられたタイミングデータである第1基準タイミングデータを受信する第1参照用データ受信手段と、

前記樂曲の音声波形の他の一部を抽象化した第2参照用データおよび前記第2参照用データに対応づけられたタイミングデータである第2基準タイミングデータ

タを受信する第2参照用データ受信手段と、

楽曲の音声波形を示す音声データを受信する音声データ受信手段と、

前記音声データの中から、前記第1参照用データが表す音声波形に類似する音声波形を示すデータを第1部分データとして選択する第1部分データ選択手段と

前記音声データの中から、前記第2参照用データが表す音声波形に類似する音声波形を示すデータを第2部分データとして選択する第2部分データ選択手段と

前記第1基準タイミングデータが示す第1基準時点を、前記第1部分データの再生タイミングに一致させ、かつ前記第2基準タイミングデータが示す第2基準時点を、前記第2部分データの再生タイミングに一致させるように、前記演奏タイミングデータにより示される前記制御データの各データの送信タイミングを調整するタイミング調整手段と、

前記タイミング調整手段により調整された送信タイミングに基づき、前記制御データの送信を行う制御データ送信手段と

を備えることを特徴とする再生装置。

【請求項3】 演奏の制御を指示する制御データおよび該演奏の制御の実行タイミングを指示する演奏タイミングデータからなる演奏データを受信する演奏データ受信過程と、

楽曲の音声波形を示す音声データを受信する音声データ受信過程と、

前記音声データから、前記楽曲の一部の音声波形を示す第1部分データを取り出す第1部分データ取得過程と、

前記第1部分データが表す音声波形を抽象化した第1参照用データを生成する第1参照用データ生成過程と、

前記音声データから、前記楽曲の一部とは異なる音声波形を示す第2部分データを取り出す第2部分データ取得過程と、

前記第2部分データが表す音声波形を抽象化した第2参照用データを生成する第2参照用データ生成過程と、

前記第1参照用データ、前記第1部分データの再生タイミングを示す第1基準

タイミングデータ、前記第2参照用データおよび前記第2部分データの再生タイミングを示す第2基準タイミングデータを、前記演奏データと共に記録する記録過程と

を備えることを特徴とする記録方法。

【請求項4】 演奏の制御を指示する制御データおよび該演奏の制御の実行タイミングを指示する演奏タイミングデータからなる演奏データを受信する演奏データ受信過程と、

楽曲の音声波形の一部を抽象化した第1参照用データおよび前記第1参照用データに対応づけられたタイミングデータである第1基準タイミングデータを受信する第1参照用データ受信過程と、

前記楽曲の音声波形の他の一部を抽象化した第2参照用データおよび前記第2参照用データに対応づけられたタイミングデータである第2基準タイミングデータを受信する第2参照用データ受信過程と、

楽曲の音声波形を示す音声データを受信する音声データ受信過程と、

前記音声データの中から、前記第1参照用データが表す音声波形に類似する音声波形を示すデータを第1部分データとして選択する第1部分データ選択過程と

前記音声データの中から、前記第2参照用データが表す音声波形に類似する音声波形を示すデータを第2部分データとして選択する第2部分データ選択過程と

前記第1基準タイミングデータが示す第1基準時点を、前記第1部分データの再生タイミングに一致させ、かつ前記第2基準タイミングデータが示す第2基準時点を、前記第2部分データの再生タイミングに一致させるように、前記演奏タイミングデータにより示される前記制御データの各データの送信タイミングを調整するタイミング調整過程と、

前記タイミング調整過程において調整された送信タイミングに基づき、前記制御データの送信を行う制御データ送信過程と

を備えることを特徴とする再生方法。

【請求項5】 演奏の制御を指示する制御データおよび該演奏の制御の実行

タイミングを指示する演奏タイミングデータからなる演奏データを受信する演奏データ受信処理と、

楽曲の音声波形を示す音声データを受信する音声データ受信処理と、

前記音声データから、前記楽曲の一部の音声波形を示す第1部分データを取り出す第1部分データ取得処理と、

前記第1部分データが表す音声波形を抽象化した第1参照用データを生成する第1参照用データ生成処理と、

前記音声データから、前記楽曲の一部とは異なる音声波形を示す第2部分データを取り出す第2部分データ取得処理と、

前記第2部分データが表す音声波形を抽象化した第2参照用データを生成する第2参照用データ生成処理と、

前記第1参照用データ、前記第1部分データの再生タイミングを示す第1基準タイミングデータ、前記第2参照用データおよび前記第2部分データの再生タイミングを示す第2基準タイミングデータを、前記演奏データと共に記録する記録処理と

をコンピュータに実行させるプログラム。

【請求項6】 演奏の制御を指示する制御データおよび該演奏の制御の実行タイミングを指示する演奏タイミングデータからなる演奏データを受信する演奏データ受信処理と、

楽曲の音声波形の一部を抽象化した第1参照用データおよび前記第1参照用データに対応づけられたタイミングデータである第1基準タイミングデータを受信する第1参照用データ受信処理と、

前記楽曲の音声波形の他の一部を抽象化した第2参照用データおよび前記第2参照用データに対応づけられたタイミングデータである第2基準タイミングデータを受信する第2参照用データ受信処理と、

楽曲の音声波形を示す音声データを受信する音声データ受信処理と、

前記音声データの中から、前記第1参照用データが表す音声波形に類似する音声波形を示すデータを第1部分データとして選択する第1部分データ選択処理と

前記音声データの中から、前記第2参照用データが表す音声波形に類似する音声波形を示すデータを第2部分データとして選択する第2部分データ選択処理と

前記第1基準タイミングデータが示す第1基準時点を、前記第1部分データの再生タイミングに一致させ、かつ前記第2基準タイミングデータが示す第2基準時点を、前記第2部分データの再生タイミングに一致させるように、前記演奏タイミングデータにより示される前記制御データの各データの送信タイミングを調整するタイミング調整処理と、

前記タイミング調整処理において調整された送信タイミングに基づき、前記制御データの送信を行う制御データ送信処理と
をコンピュータに実行させるプログラム。

【請求項7】 演奏の制御を指示する制御データおよび該演奏の制御の実行タイミングを指示する演奏タイミングデータからなる演奏データと、

楽曲の音声波形の一部を抽象化した第1参照用データおよび前記第1参照用データに対応づけられたタイミングデータである第1タイミングデータと、

前記楽曲の音声波形の他の一部を抽象化した第2参照用データおよび前記第2参照用データに対応づけられたタイミングデータである第2タイミングデータ
が記録された記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、オーディオデータの再生に同期して、楽曲の演奏制御に関する情報を含む演奏データの再生を行う装置および方法に関する。

【0002】

【従来の技術】

楽曲を再生するための手段として、音楽CD(Compact Disc)などの記憶媒体から音声データを読み出し、読み出された音声データから音声を生成して出力する装置がある。また、楽曲を再生するための他の手段として、FD(Floppy Disk)などの記憶媒体から楽曲の演奏制御に関する情報を

含むデータを読み出し、読み出されたデータを用いて音源装置の発音を制御することにより自動演奏を行う装置がある。楽曲の演奏制御に関する情報を含むデータとしては、MIDI (Musical Instrument Digital Interface) 規格に従って作成されたMIDIデータがある。

【0003】

もし、一般に流通している音楽CDに記録されている音声データの再生に対し、MIDIデータによる自動演奏を同期させることができれば、音楽CDから再生される楽曲に対し、MIDIデータにより好みの伴奏を付加でき、便利である。

【0004】

一般に流通している音楽CDとMIDIデータによる演奏の同期に関し、応用が可能と思われる従来技術として、演奏に付与する照明、画像、音響等の効果の内容を予め演奏データに記述しておき、演奏により発生するイベントに対応づけて演奏データに記録されたフラグを用いて各種効果の制御を正しいタイミングで行うものがある（例えば、特許文献1参照）。また、MIDIデータの同期制御に関する従来技術として、非同期に発生するMIDIイベント間に同期用データを補充するものがある（例えば、特許文献2参照）。

【0005】

【特許文献1】

特開2001-195061号公報

【特許文献2】

特願2001-215958号公報

【0006】

【発明が解決しようとする課題】

従来の技術によれば、MIDIデータに対し、正しいタイミングで音楽CDの音声データの再生を開始することができる。しかしながら、単に音楽CDの音声データの再生を正しいタイミングで開始するだけでは、MIDIデータの記録に用いられたクロック速度とMIDIデータの再生に用いられるクロック速度とがわずかに異なる等の理由で、音楽CDの再生による楽曲とMIDIデータの再生

による演奏とが、曲が進むにつれずれてしまうという問題があった。

【0007】

さらに、同じ楽曲に関し異なる版の音楽CDが存在する場合、それぞれの版によって音声データの再生開始のタイミングから楽曲が実際に始まるタイミングまでの沈黙時間の長さが異なっていることがある。さらに、異なる版の音楽CDの音声データは、記録に用いられたクロック速度がそれぞれ異なっていることがある。このため、それらを同じ再生装置により再生したとしても、楽曲の速度がわずかに異なることがある。このように沈黙時間や速度に差が生じるのは、多くの音楽CDに記録されている音声データが、実際の演奏を記録したマスターデータに対し、音響効果等の編集作業を施した後に再度記録を行ったものであるためである。すなわち、その編集作業により沈黙時間の一部がカットされたり、編集作業後の音声データを再度記録する際に用いられるクロック速度が正確に同じではなかったりするため、版により沈黙時間や速度が異なる結果となる。

【0008】

上述のように楽曲に関して複数の版の音楽CDがある場合には、それぞれの版に対して異なる再生開始のタイミングを設定する必要がある。また、たとえMIDIデータの記録時および再生時のクロック速度等にずれが無くとも、版によってMIDIデータに対し音声データの楽曲が速く進行したり遅く進行したりする。従って、同じ楽曲であるにもかかわらず、それぞれの版に対し異なるMIDIデータを準備する必要があり、不便であった。

【0009】

上述した状況に鑑み、本発明は、同じ楽曲に関する異なる複数の版の音声データに対し、同期再生が可能なMIDIデータ等の演奏データの記録装置、再生装置、記録方法、再生方法、プログラムおよび記録媒体を提供することを目的とする。

【0010】

【課題を解決するための手段】

以上説明した課題を解決するため、本発明は、演奏の制御を指示する制御データおよび該演奏の制御の実行タイミングを指示する演奏タイミングデータからな

る演奏データを受信する演奏データ受信手段と、楽曲の音声波形を示す音声データを受信する音声データ受信手段と、前記音声データから、前記楽曲の一部の音声波形を示す第1部分データを取り出す第1部分データ取得手段と、前記第1部分データが表す音声波形を抽象化した第1参照用データを生成する第1参照用データ生成手段と、前記音声データから、前記楽曲の一部とは異なる音声波形を示す第2部分データを取り出す第2部分データ取得手段と、前記第2部分データが表す音声波形を抽象化した第2参照用データを生成する第2参照用データ生成手段と、前記第1参照用データ、前記第1部分データの再生タイミングを示す第1基準タイミングデータ、前記第2参照用データおよび前記第2部分データの再生タイミングを示す第2基準タイミングデータを、前記演奏データと共に記録する記録手段とを備えることを特徴とする記録装置を提供する。

【0011】

また好ましい態様において、本発明にかかる記録装置は、計時を行う計時手段をさらに有し、前記記録手段は、前記音声データ受信手段が前記第1部分データを受信する時点における前記計時の結果に基づき生成される前記第1基準タイミングデータと、前記音声データ受信手段が前記第2部分データを受信する時点における前記計時の結果に基づき生成される前記第2基準タイミングデータを記録することを特徴としている。

【0012】

また別の好ましい態様において、本発明にかかる記録装置は、前記音声データの再生タイミングを示すタイムコードを受信するタイムコード受信手段をさらに有し、前記記録手段は、前記音声データ受信手段が前記第1部分データを受信する時点に前記タイムコード受信手段が受信するタイムコードに基づき生成される前記第1基準タイミングデータと、前記音声データ受信手段が前記第2部分データを受信する時点に前記タイムコード受信手段が受信するタイムコードに基づき生成される前記第2基準タイミングデータとを記録することを特徴としている。

【0013】

また別の好ましい態様において、本発明にかかる記録装置は、前記音声データの所定期間の変化傾向を示す第1指標を生成する第1指標生成手段と、前記音声

データの前記所定期間より長い期間の変化傾向を示す第2指標を生成する第2指標生成手段と、前記第1指標と前記第2指標とを比較して、前記音声データの急激な変化を検出する検出手段とをさらに有し、前記記録手段は、前記急激な変化のタイミングを示す変化タイミングデータをさらに記録する機能を有することを特徴としている。

【0014】

また、本発明は、演奏の制御を指示する制御データおよび該演奏の制御の実行タイミングを指示する演奏タイミングデータからなる演奏データを受信する演奏データ受信手段と、楽曲の音声波形の一部を抽象化した第1参照用データおよび前記第1参照用データに対応づけられたタイミングデータである第1基準タイミングデータを受信する第1参照用データ受信手段と、前記楽曲の音声波形の他の一部を抽象化した第2参照用データおよび前記第2参照用データに対応づけられたタイミングデータである第2基準タイミングデータを受信する第2参照用データ受信手段と、楽曲の音声波形を示す音声データを受信する音声データ受信手段と、前記音声データの中から、前記第1参照用データが表す音声波形に類似する音声波形を示すデータを第1部分データとして選択する第1部分データ選択手段と、前記音声データの中から、前記第2参照用データが表す音声波形に類似する音声波形を示すデータを第2部分データとして選択する第2部分データ選択手段と、前記第1基準タイミングデータが示す第1基準時点を、前記第1部分データの再生タイミングに一致させ、かつ前記第2基準タイミングデータが示す第2基準時点を、前記第2部分データの再生タイミングに一致させるように、前記演奏タイミングデータにより示される前記制御データの各データの送信タイミングを調整するタイミング調整手段と、前記タイミング調整手段により調整された送信タイミングに基づき、前記制御データの送信を行う制御データ送信手段とを備えることを特徴とする再生装置を提供する。

【0015】

また好ましい態様において、本発明にかかる再生装置は、計時を行う計時手段をさらに有し、前記タイミング調整手段は、前記第1基準時点を、前記音声データ受信手段が前記第1部分データを受信する時点における前記計時の結果に基づ

く前記第1部分データの再生タイミングに一致させ、かつ前記第2基準時点を、前記音声データ受信手段が前記第2部分データを受信する時点における前記計時の結果に基づく前記第2部分データの再生タイミングに一致させるように、前記演奏タイミングデータにより示される前記制御データの各データの送信タイミングを調整することを特徴としている。

【0016】

また別の好ましい態様において、本発明にかかる再生装置は、前記音声データの再生タイミングを示すタイムコードを受信するタイムコード受信手段をさらに有し、前記タイミング調整手段は、前記第1基準時点を、前記音声データ受信手段が前記第1部分データを受信する時点に前記タイムコード受信手段が受信するタイムコードに基づく前記第1部分データの再生タイミングに一致させ、かつ前記第2基準時点を、前記音声データ受信手段が前記第2部分データを受信する時点に前記タイムコード受信手段が受信するタイムコードに基づく前記第2部分データの再生タイミングに一致させるように、前記演奏タイミングデータにより示される前記制御データの各データの送信タイミングを調整することを特徴としている。

【0017】

また別の好ましい態様において、本発明にかかる再生装置は、調整データを受信する調整データ受信手段をさらに有し、前記第1部分データ選択手段が前記第1部分データを選択できない場合、もしくは前記第2部分データ選択手段が前記第2部分データを選択できない場合、前記タイミング調整手段は、前記第1基準タイミングデータが示す第1基準時点もしくは前記第2基準タイミングデータが示す第2基準時点を、前記調整データが示すタイミングに一致させるように、前記演奏タイミングデータにより示される前記制御データの各データの送信タイミングを調整することを特徴としている。

【0018】

また別の好ましい態様において、本発明にかかる再生装置は、楽曲の音声波形の急激な変化のタイミングを示す第1変化タイミングデータを受信する変化タイミングデータ受信手段と、前記音声データの所定期間の変化傾向を示す第1指標

を生成する第1指標生成手段と、前記音声データの前記所定期間より長い期間の変化傾向を示す第2指標を生成する第2指標生成手段と、前記第1指標と前記第2指標とを比較して、前記音声データの急激な変化を検出し、前記急激な変化のタイミングを示す第2変化タイミングデータを生成する変化タイミングデータ生成手段と、前記第1変化タイミングデータおよび前記第2変化タイミングデータが示すタイミングに基づき、第1調整データおよび第2調整データを生成する調整データ生成手段とをさらに有し、前記第1部分データ選択手段が前記第1部分データを選択できない場合、もしくは前記第2部分データ選択手段が前記第2部分データを選択できない場合、前記タイミング調整手段は、前記第1基準タイミングデータが示す第1基準時点を前記第1調整データが示すタイミングに一致させ、かつ前記第2基準タイミングデータが示す第2基準時点を、前記第2調整データが示すタイミングに一致させるように、前記演奏タイミングデータにより示される前記制御データの各データの送信タイミングを調整することを特徴としている。

【0019】

また、本発明は、演奏の制御を指示する制御データおよび該演奏の制御の実行タイミングを指示する演奏タイミングデータからなる演奏データを受信する演奏データ受信過程と、楽曲の音声波形を示す音声データを受信する音声データ受信過程と、前記音声データから、前記楽曲の一部の音声波形を示す第1部分データを取り出す第1部分データ取得過程と、前記第1部分データが表す音声波形を抽象化した第1参照用データを生成する第1参照用データ生成過程と、前記音声データから、前記楽曲の一部とは異なる音声波形を示す第2部分データを取り出す第2部分データ取得過程と、前記第2部分データが表す音声波形を抽象化した第2参照用データを生成する第2参照用データ生成過程と、前記第1参照用データ、前記第1部分データの再生タイミングを示す第1基準タイミングデータ、前記第2参照用データおよび前記第2部分データの再生タイミングを示す第2基準タイミングデータを、前記演奏データと共に記録する記録過程とを備えることを特徴とする記録方法を提供する。

【0020】

また、本発明は、演奏の制御を指示する制御データおよび該演奏の制御の実行タイミングを指示する演奏タイミングデータからなる演奏データを受信する演奏データ受信過程と、楽曲の音声波形の一部を抽象化した第1参照用データおよび前記第1参照用データに対応づけられたタイミングデータである第1基準タイミングデータを受信する第1参照用データ受信過程と、前記楽曲の音声波形の他の一部を抽象化した第2参照用データおよび前記第2参照用データに対応づけられたタイミングデータである第2基準タイミングデータを受信する第2参照用データ受信過程と、楽曲の音声波形を示す音声データを受信する音声データ受信過程と、前記音声データの中から、前記第1参照用データが表す音声波形に類似する音声波形を示すデータを第1部分データとして選択する第1部分データ選択過程と、前記音声データの中から、前記第2参照用データが表す音声波形に類似する音声波形を示すデータを第2部分データとして選択する第2部分データ選択過程と、前記第1基準タイミングデータが示す第1基準時点を、前記第1部分データの再生タイミングに一致させ、かつ前記第2基準タイミングデータが示す第2基準時点を、前記第2部分データの再生タイミングに一致させるように、前記演奏タイミングデータにより示される前記制御データの各データの送信タイミングを調整するタイミング調整過程と、前記タイミング調整過程において調整された送信タイミングに基づき、前記制御データの送信を行う制御データ送信過程とを備えることを特徴とする再生方法を提供する。

【0021】

また、本発明は、演奏の制御を指示する制御データおよび該演奏の制御の実行タイミングを指示する演奏タイミングデータからなる演奏データを受信する演奏データ受信処理と、楽曲の音声波形を示す音声データを受信する音声データ受信処理と、前記音声データから、前記楽曲の一部の音声波形を示す第1部分データを取り出す第1部分データ取得処理と、前記第1部分データが表す音声波形を抽象化した第1参照用データを生成する第1参照用データ生成処理と、前記音声データから、前記楽曲の一部とは異なる音声波形を示す第2部分データを取り出す第2部分データ取得処理と、前記第2部分データが表す音声波形を抽象化した第2参照用データを生成する第2参照用データ生成処理と、前記第1参照用データ

、前記第1部分データの再生タイミングを示す第1基準タイミングデータ、前記第2参照用データおよび前記第2部分データの再生タイミングを示す第2基準タイミングデータを、前記演奏データと共に記録する記録処理とをコンピュータに実行させるプログラムを提供する。

【0022】

また、本発明は、演奏の制御を指示する制御データおよび該演奏の制御の実行タイミングを指示する演奏タイミングデータからなる演奏データを受信する演奏データ受信処理と、楽曲の音声波形の一部を抽象化した第1参照用データおよび前記第1参照用データに対応づけられたタイミングデータである第1基準タイミングデータを受信する第1参照用データ受信処理と、前記楽曲の音声波形の他の一部を抽象化した第2参照用データおよび前記第2参照用データに対応づけられたタイミングデータである第2基準タイミングデータを受信する第2参照用データ受信処理と、楽曲の音声波形を示す音声データを受信する音声データ受信処理と、前記音声データの中から、前記第1参照用データが表す音声波形に類似する音声波形を示すデータを第1部分データとして選択する第1部分データ選択処理と、前記音声データの中から、前記第2参照用データが表す音声波形に類似する音声波形を示すデータを第2部分データとして選択する第2部分データ選択処理と、前記第1基準タイミングデータが示す第1基準時点を、前記第1部分データの再生タイミングに一致させ、かつ前記第2基準タイミングデータが示す第2基準時点を、前記第2部分データの再生タイミングに一致させるように、前記演奏タイミングデータにより示される前記制御データの各データの送信タイミングを調整するタイミング調整処理と、前記タイミング調整処理において調整された送信タイミングに基づき、前記制御データの送信を行う制御データ送信処理とをコンピュータに実行させるプログラムを提供する。

【0023】

また、本発明は、演奏の制御を指示する制御データおよび該演奏の制御の実行タイミングを指示する演奏タイミングデータからなる演奏データと、楽曲の音声波形の一部を抽象化した第1参照用データおよび前記第1参照用データに対応づけられたタイミングデータである第1タイミングデータと、前記楽曲の音声波形

の他的一部を抽象化した第2参照用データおよび前記第2参照用データに対応づけられたタイミングデータである第2タイミングデータが記録された記録媒体を提供する。

【0024】

かかる構成による装置、方法、プログラム、および記録媒体を用いると、音声データを再生する場合、音声データが表す波形の類似性により、参照用データの音声データに対する時間的な位置を決定することができ、さらに参照用データの時間的な位置に基づき、制御データの再生のタイミングを決定することができる。その結果、音声データと制御データとの同期再生が実現される。

【0025】

【発明の実施の形態】

[1：実施形態]

以下、本発明の好適な実施形態として、音楽CDに記録された音声データの再生と、FDに記録されたMIDIデータによる演奏との同期再生を実現する装置およびその装置の動作を説明する。なお、本発明において利用可能な音声データは音楽CDに記録された音声データに限らず、あらゆる形態の音声データが利用可能である。さらに、本発明において利用可能な演奏データはMIDIデータに限らず、あらゆる形態の演奏データが利用可能である。

【0026】

[1. 1：構成、機能、およびデータフォーマット]

[1. 1. 1：全体構成]

図1は、本発明の実施形態に係る同期記録再生装置SSの構成を示す図である。同期記録再生装置SSは、CDドライブ1、FDドライブ2、自動演奏ピアノ部3、発音部4、操作表示部5およびコントローラ部6により構成されている。CDドライブ1、FDドライブ2、自動演奏ピアノ部3、発音部4および操作表示部5はそれぞれ通信線によりコントローラ部6と接続されている。また、自動演奏ピアノ部3と発音部4は、通信線により直接接続されている。

【0027】

[1. 1. 2：CDドライブ]

音楽CDに記録されているオーディオデータは、音声情報を示す音声データ、音声データの再生のタイミングを示すタイムコード、および各音声データの開始時間等の目次情報を含んでいる。CDドライブ1は、コントローラ部6からの指示に従って音声データを音楽CDから読み出し、読み出した音声データを順次出力する装置である。CDドライブ1は通信線により、コントローラ部6の通信インターフェース65と接続されている。

【0028】

CDドライブ1から出力される音声データは、サンプリング周波数44100Hz、量子化ビット数16の左右2チャンネルからなるデジタル音声データである。なお、CDドライブ1から出力されるデータには、タイムコードは含まれていない。CDドライブ1の構成は、音声データをデジタル出力可能な一般的なCDドライブと同様であるため、その説明は省略する。

【0029】

[1. 1. 3 : FDドライブ]

FDドライブ2は、SMF (Standard MIDI File) をFDに記録し、またFDに記録されているSMFを読み取り、読み取ったSMFを送信する装置である。FDドライブ2は通信線により、コントローラ部6の通信インターフェース65と接続されている。なお、FDドライブ2の構成は、一般的なFDドライブと同様であるため、その説明は省略する。

【0030】

[1. 1. 4 : イベントデータおよびSMF]

SMFは、MIDI規格に従った演奏制御の内容を示すデータであるイベントデータと、各イベントデータの実行タイミングを示すデータであるデルタタイムを含むファイルである。図2および図3を用いて、イベントデータおよびSMFのフォーマットを説明する。

【0031】

図2にはSMFのフォーマットの概略が示されている。SMFはヘッダチャンクとトラックチャンクから成る。ヘッダチャンクには、トラックチャンネルに含まれるデータのフォーマットや時間単位等に関する制御データが含まれている。

トラックチャンクには、イベントデータと、各イベントデータの実行タイミングを示すデルタタイムが含まれている。以下の説明においては、デルタタイムは、基準となる時点からの絶対的な時間とし、その単位を秒で表すこととするが、実際には、デルタタイムはMIDI規格に従った表現形式により扱われる。また、デルタタイムは相対時間で表すようにしてもよいし、別の表現でもよい。

【0032】

図3にはSMF中のイベントデータの例として、ノートオンイベント、ノートオフイベント、およびシステムエクスクルーシブイベントが示されている。以下、システムエクスクルーシブイベント以外のイベントデータをシステムエクスクルーシブイベントと区別する目的で、「演奏イベント」と呼ぶ。イベントデータは時間情報を持たず、リアルタイムに楽音の発音、消音、およびその他の制御を行う目的で利用される。

なお、本明細書においては、MIDIデータとはMIDI規格に従って作成されるデータの総称である。

【0033】

[1. 1. 5：自動演奏ピアノ部]

自動演奏ピアノ部3は、同期記録再生装置SSのユーザによる鍵操作およびペダル操作に応じて、音響的なピアノ音および電子的な楽音合成によるピアノ音を出力する楽音発生装置である。また、自動演奏ピアノ部3はユーザによる鍵操作およびペダル操作に応じて演奏イベントを生成し、生成した演奏イベントを送信する。さらに、自動演奏ピアノ部3は演奏イベントを受信して、受信した演奏イベントに応じて、音響的なピアノ音および電子的な楽音合成によるピアノ音による自動演奏を行う。

【0034】

自動演奏ピアノ部3は、ピアノ31、キーセンサ32と、ペダルセンサ33、MIDIイベント制御回路34、音源35、および駆動部36から構成されている。

【0035】

キーセンサ32およびペダルセンサ33は、それぞれピアノ31の複数の鍵お

より複数のペダルの各々に配設され、それぞれ鍵およびペダルの位置を検出する。キーセンサ32およびペダルセンサ33は、検出された位置情報を、それぞれの鍵およびペダルに対応した識別番号と共に、MIDIイベント制御回路34に送信する。

【0036】

MIDIイベント制御回路34は、キーセンサ32およびペダルセンサ33から、それぞれの鍵およびペダルの位置情報を、鍵およびペダルの識別情報、および時間情報と共に受信し、これらの情報から即時にノートオンイベントやノートオフイベント等の演奏イベントを生成し、生成した演奏イベントをコントローラ部6および音源35に出力する回路である。さらに、MIDIイベント制御回路34は、コントローラ部6から演奏イベントを受信し、受信した演奏イベントを音源35もしくは駆動部36に転送する機能も持つ。なお、MIDIイベント制御回路34が、コントローラ部6から受信する演奏イベントを音源35と駆動部36のいずれに転送するかは、コントローラ部6の指示による。

【0037】

音源35は、MIDIイベント制御回路34から演奏イベントを受信し、受信した演奏イベントに基づいて、各種楽器の音情報を左右2チャネルのデジタル音声データとして出力する装置である。音源35は受信した演奏イベントによって指示された音高のデジタル音声データを電子的に合成し、発音部4のミキサ41に送信する。

【0038】

駆動部36は、ピアノ31の各鍵および各ペダルに配設され、それらを駆動するソレノイド群およびそれらのソレノイド群を制御する制御回路から構成される。駆動部36の制御回路は、MIDIイベント制御回路から演奏イベントを受信すると、対応する鍵もしくはペダルに配設されたソレノイドへ供給する電流量を調節し、ソレノイドが発生する磁力を制御することにより、演奏イベントに応じた鍵もしくはペダルの動作を実現する。

【0039】

[1. 1. 6 : 発音部]

発音部4は、自動演奏ピアノ部3およびコントローラ部6から音声データを受信し、受信した音声データを音に変換して出力する装置である。発音部4は、ミキサ41、D/Aコンバータ42、アンプ43、およびスピーカ44から構成されている。

【0040】

ミキサ41は左右2チャンネルからなるデジタル音声データを複数受信し、それらを左右1組のデジタル音声データに変換するデジタルステレオミキサである。ミキサ41は自動演奏ピアノ部3の音源35からデジタル音声データを受信すると同時に、CDドライブ1により音楽CDから読み出されたデジタル音声データを、コントローラ部6を介して受信する。ミキサ41は受信したこれらのデジタル音声データを算術平均し、左右1組のデジタル音声データとしてD/Aコンバータ42に送信する。

【0041】

D/Aコンバータ42は、ミキサ41からデジタル音声データを受信し、受信したデジタル音声データをアナログ音声信号に変換し、アンプ43に出力する。アンプ43は、D/Aコンバータ42から入力されるアナログ音声信号を増幅し、スピーカ44に出力する。スピーカ44は、アンプ43から入力される増幅されたアナログ音声信号を音に変換する。その結果、音楽CDに記録された音声データおよび音源35が生成する音声データは、ステレオの音として発音部4から出力される。

【0042】

[1. 1. 7 : 操作表示部]

操作表示部5は、同期記録再生装置SSのユーザが同期記録再生装置SSの各種操作を行う際に用いるユーザインタフェースである。操作表示部5はユーザが同期記録再生装置SSに指示を与える際に押下するキーパッド、ユーザが同期記録再生装置SSの状態を確認するための液晶ディスプレイ等を有している。ユーザによりキーパッドが押下されると、操作表示部5は押下されたキーパッドに対応する信号をコントローラ部6に出力する。また、操作表示部5はコントローラ部6から文字や図形の情報を含むビットマップデータを受信すると、受信したビ

ットマップデータに基づき液晶ディスプレイに文字や図形を表示する。

【0043】

[1. 1. 8 : コントローラ部]

コントローラ部6は、同期記録再生装置SSの全体を制御する装置であり、その構成は一般的なコンピュータ本体と同様である。すなわち、コントローラ部6はROM(Read Only Memory)61、CPU(Central Processing Unit)62、DSP(Digital Signal Processor)63、RAM(Random Access Memory)64、および通信インターフェース65から構成されている。これらの構成要素はバスにより相互に接続されている。また、コントローラ部6は図示せぬクロックを有し、コントローラ部6の各構成部の動作はこのクロックにより生成されるクロック信号に基づいて同期が取られている。

【0044】

ROM61は各種の制御用プログラムを格納する不揮発性のメモリである。ROM61が格納する制御用プログラムには、一般的な制御処理を行うプログラムに加え、後述するSMFの記録動作および再生動作における処理をCPU62に実行させるプログラムが含まれている。CPU62は汎用的な処理を実行可能なマイクロプロセッサであり、ROM61から制御用プログラムを読み込み、読み込んだ制御用プログラムに従った制御処理を行う。DSP63はデジタル音声データを高速に処理可能なマイクロプロセッサであり、CPU62の制御に従い、CDドライブ1やFDドライブ2からコントローラ部6が受信するデジタル音声データに対し、後述する相関判定用データ生成処理および相関判定処理において必要とされるフィルタ処理等の処理を施し、その結果得られるデータをCPU62に送信する。RAM64は揮発性メモリであり、CPU62およびDSP63が利用するデータを一時的に格納する。通信インターフェース65は各種フォーマットのデジタルデータを送受信可能なインターフェースであり、CDドライブ1、FDドライブ2、自動演奏ピアノ部3、発音部4、および操作表示部5との間で送受信されるデジタルデータに対し必要なフォーマット変換を行い、それらの各装置とコントローラ部6との間のデータの中継を行う。

【0045】

[1. 2 : 動作]

続いて、同期記録再生装置SSの動作を説明する。

[1. 2. 1 : 記録動作]

まず、同期記録再生装置SSのユーザが市販の音楽CDの再生に合わせてピアノを演奏し、その演奏の情報をSMFとしてFDに記録する際の同期記録再生装置SSにおける動作を説明する。なお、以下に説明する記録動作において用いられる音楽CDを、後述する再生動作において用いられる音楽CDと区別するために、音楽CD-Aと呼ぶ。さらに、音楽CD-Aには複数の楽曲の音声データが記録されているが、以下の説明においてはその中の1曲である楽曲Nに関する音声データNAについて、記録動作が行われるものとする。

【0046】

[1. 2. 1. 1 : 記録の開始操作]

ユーザは、音楽CD-AをCDドライブ1に、また十分な空き容量を持つFDをFDドライブ2にセットする。続いて、ユーザは音楽CD-Aに含まれる楽曲Nに対する演奏データの記録開始を、操作表示部5のキーパッドを押下することにより指示する。操作表示部5は押下されたキーパッドに対応する信号をコントローラ部6に出力する。

【0047】

コントローラ部6のCPU62は操作表示部5から演奏データの記録開始に対応する信号を受信すると、CDドライブ1に音声データNAの再生命令を送信する。この再生命令に応じて、CDドライブ1は音楽CD-Aから音声データNAを読み出し、読み出した音声データNAをコントローラ部6に順次送信する。

【0048】

コントローラ部6はCDドライブ1から、1/44100秒ごとに左右1組のデータを受信する。以下、左右1組のデータの値を(R(n), L(n))と表し、この1組のデータの値、もしくは後述する相関判定用データ生成処理もしくは管理用イベント生成処理においてこの1組のデータから生成される各データの値を「サンプル値」と呼ぶ。また、異なるサンプル値を区別する場合には、サン

プル値 (n) のように表す。n は音声データの順序を表す整数で、先頭のデータから順に 0、1、2、…と増加する。R (n) および L (n) はそれぞれ右チャンネルのデータおよび左チャンネルのデータの値を示し、-32768～32767 のいずれかの整数である。

【0049】

[1. 2. 1. 2 : 音声データの発音部への送信]

CPU62 は、CD ドライブ 1 から音声データ NA の最初のサンプル値、すなわちサンプル値 (0) を受信すると、クロックから得られるクロック信号に基づき、その時点（以下、「基準時点 P」と呼ぶ）を基準とする計時を開始する。すなわち、基準時点 P における時間は 0.00 秒である。

【0050】

CPU62 はサンプル値 (0) に続き、サンプル値 (1)、サンプル値 (2)、…を順次受信する。CPU62 は受信したサンプル値を順次、発音部 4 に送信する。発音部 4 はサンプル値を受信すると、これを音に変換し出力する。その結果、ユーザは楽曲 N を聞くことができる。なお、CPU62 は楽曲 N の最後のサンプル値を発音部 4 に送信するまで、この送信処理を続ける。

【0051】

[1. 2. 1. 3 : 音声データのキューへの記録]

CPU62 は受信したサンプル値を順次、発音部 4 に送信すると同時に、受信したサンプル値を、サンプル値の受信時刻を示す時間情報と共に、RAM64 のキューに記録する。この時間情報は、基準時点 P からの経過時間である。以下、サンプル値 (n) に対応する時間情報を時間情報 (n) と呼ぶ。

【0052】

本実施形態においては、例として、このキューに記録可能なサンプル値の数は 1323000 組とし、CPU62 はキューに記録されているサンプル値の数が 1323000 組に達した後は、新たにサンプル値を受信すると、キューの先頭のサンプル値を削除し、キューの末尾に新たに受信したサンプル値を記録する。なお、キューに記録される 1323000 組のサンプル値は 30 秒間分の音声データである。CPU62 は音声データ NA の最後のサンプル値を受信するまで、

サンプル値および時間情報をキューに記録する処理を続ける。

【0053】

【1. 2. 1. 4 : 先頭参照用未加工音声データの記録】

CPU62は、上述したサンプル値の発音部4への送信およびキューへの記録を継続する一方で、音声データにおいて実際に楽曲が始まる時点以降の一定時間に対応するサンプル値を別途、RAM64に記録する処理を行う。本実施形態においては、例として、この処理で記録されるサンプル値は 2^{16} 組、すなわち65536組とする。なお、65536組のサンプル値は約1.49秒間分の音声データである。以下、この65536組のサンプル値を「先頭参照用未加工音声データ」と呼ぶ。続いて、先頭参照用未加工音声データの記録処理を説明する。

【0054】

まず、CPU62は受信した各サンプル値について、各サンプル値の左右いずれかの絶対値が予め定められた閾値を超えているか否かを、順次判定する。以下、この判定処理を「閾値判定処理」と呼ぶ。本実施形態においては、例として、その閾値を1000とする。従って、R(n)およびL(n)のいずれかの絶対値が1000より大きい場合、CPU62は閾値判定処理で肯定的な結果を得る。

【0055】

例えば、音声データNAに関し、サンプル値(50760)において初めて、これに対応したデータであるR(50760)もしくはL(50760)の絶対値が1000を超えるものとする。この場合、サンプル値(0)～サンプル値(50759)に対するCPU62の閾値判定処理は否定的な結果を得る。これは音声データNAの冒頭に約1.15秒間の沈黙もしくは沈黙に近い部分が存在することを示す。

【0056】

CPU62は受信したサンプル値に関する閾値判定処理について肯定的な結果を得ると、それ以降に受信するサンプル値に対する閾値判定処理は行わず、基準時点Pを基準としたそのサンプル値の受信時刻を示す時間情報をRAM64に記録する。この時間情報を以下、「先頭時間情報」と呼ぶ。さらに、CPU62は

閾値判定処理の結果が肯定的となったサンプル値以降に受信する65536組のサンプル値を、先頭参照用未加工音声データとしてRAM64に記録する。

【0057】

例えば、サンプル値(50760)において初めて閾値判定処理について肯定的な結果が得られた場合、先頭参照用未加工音声データはサンプル値(50760)～サンプル値(116295)となり、また先頭時間情報は、サンプル値(50760)の受信時刻を示す時間情報、すなわち約1.15秒となる。

【0058】

[1. 2. 1. 5: 先頭参照用加工済音声データの生成]

CPU62は、先頭参照用未加工音声データのRAM64への記録を終了すると、DSP63に、先頭参照用未加工音声データに対する相関判定用データ生成処理の実行命令を送信する。相関判定用データ生成処理とは、サンプリング周波数44100Hzの音声データから、相関判定処理に用いるためのサンプリング周波数約172.27Hzの音声データを生成する処理である。相関判定処理とは、2組の音声データの類似性を判定する処理であり、詳細は後述する。以下、図4を参照しながら、相関判定用データ生成処理を説明する。

【0059】

DSP63はCPU62より先頭参照用未加工音声データに対する相関判定用データ生成処理の実行命令を受信すると、RAM64から先頭参照用未加工音声データを読み出す(ステップS1)。続いて、DSP63は先頭参照用未加工音声データの各サンプル値の左右の値を相加平均することにより、ステレオのデータをモノラルのデータに変換する(ステップS2)。このモノラルへの変換処理は、このステップより後の処理のDSP63への負荷を軽減するための処理である。

【0060】

続いて、DSP63はモノラルに変換された一連のサンプル値に対し、ハイパスフィルタ処理を行う(ステップS3)。このハイパスフィルタ処理により、一連のサンプル値が示す音声波形における直流成分が取り除かれ、サンプル値が正負の両側に均等に分布するようになる。相関判定処理においては2組の音声データ

タの相互相關値を用いた比較判定が行われるが、相互相關値の比較の際、サンプル値は正負の両側に均等に分布すると判定の精度が高まる。すなわち、このステップにおける処理は、相関判定処理における判定の精度向上を目的とする処理である。

【0061】

続いて、DSP63はハイパスフィルタ処理を施された各サンプル値について、絶対値を求める（ステップS4）。このステップにおける処理は、各サンプル値のパワーの代替値を求めるための処理である。絶対値はパワーを示す二乗値よりも値が小さく処理が容易であるため、本実施形態においては、各サンプル値の二乗値の代替値として絶対値が用いられる。従って、DSP63の処理能力が高い場合、このステップにおいて各サンプル値の絶対値の代わりに二乗値を算出してもよい。

【0062】

続いて、DSP63はステップS4において絶対値に変換された一連のサンプル値に対し、樹形フィルタを用いたフィルタ処理を行う（ステップS5）。このステップにおける処理は、一連のサンプル値が示す音声波形から、波形の変化を捉えやすい低周波数成分を取り出すための処理である。低周波数成分を取り出すためには、ローパスフィルタが通常用いられるが、樹形フィルタはローパスフィルタと比較して、通常、DSP63への負荷が小さいため、本実施形態においてはローパスフィルタが樹形フィルタで代用されている。

【0063】

図5は、ステップS5において利用可能な樹形フィルタの一例について、その構成を示している。図5において、四角形で示される処理は遅延処理を示し、 z^{-k} におけるkは、その遅延処理における遅延時間が（サンプリング周期×k）であることを意味する。前述のとおり、音楽CDのサンプリング周波数は44100Hzであるため、サンプリング周期は1/44100秒である。一方、三角形で示される処理は乗算処理を示し、三角形の中に示される値が乗算の係数を示す。図5において、Kは次の式（1）で表される。

【数1】

$$K = \frac{44100 - \pi \times f}{44100 + \pi \times f} \dots (1)$$

【0064】

このKを係数とする乗算により、この樹形フィルタは周波数fのハイパスフィルタの機能を持つ。その結果、このステップにおけるフィルタ処理によって、一連のサンプル値が示す音声波形における直流成分が再度、取り除かれる。なお、kおよびfの値は任意に変更が可能であり、相関判定処理における判定精度が高くなるよう、経験的に求められる。

【0065】

続いて、DSP63は図4におけるステップS5においてフィルタ処理を施された一連のサンプル値に対し、さらにローパスフィルタを用いたフィルタ処理を行う（ステップS6）。このステップにおける処理は、次のステップS7において行われるダウンサンプリング処理の結果、折り返し雑音が発生することを防止するための処理である。ステップS7においては44100Hzのサンプリング周波数のデータを約172.27Hzのサンプリング周波数にダウンサンプリングするため、折り返し雑音の発生を防ぐには、その半分である約86.13Hz以上の周波数成分を取り除く必要がある。しかしながら、ステップS5における樹形フィルタを用いたフィルタ処理においては、樹形フィルタの特性により高周波数成分が十分に取り除かれない。従って、このステップにおいてローパスフィルタを用いたフィルタ処理を行うことにより、残存している約86.13Hz以上の周波数成分を取り除く。なお、DSP63の処理能力が高い場合、ステップS5とステップS6における2つのフィルタを用いたフィルタ処理の代わりに、精度の高いローパスフィルタを1つのみ用いたフィルタ処理を行ってもよい。

【0066】

続いて、DSP63はステップS6においてフィルタ処理を施された一連のサンプル値に対し、1/256のダウンサンプリング処理を行う（ステップS7）。すなわち、DSP63は、256のサンプル値ごとにサンプル値を1つ抜き取

る処理を行う。その結果、一連のサンプル値のデータ数は65536から256に減少する。以下、ステップS7の処理によって得られる一連の256のサンプル値を「先頭参照用加工済音声データ」と呼ぶ。DSP63は、先頭参照用加工済音声データをRAM64に記録する（ステップS8）。なお、ステップS7における処理は、後述する先頭参照用加工済音声データを利用した処理におけるDSP63への負荷の軽減、および先頭参照用加工済音声データをFD等に記録する際の記録容量の節減を目的とする処理である。従って、DSP63の処理能力が高い場合、もしくは利用可能な記録容量が十分である場合等はステップS7の処理を行わず、44100Hzの音声データを示すサンプル値そのものを先頭参照用加工済音声データとしてもよい。

【0067】

【1. 2. 1. 6：管理用イベントの生成】

CPU62は先頭参照用未加工音声データの先頭のサンプル値を受信すると、上述した先頭参照用未加工音声データのRAM64への記録を行うと同時に、DSP63に対し、管理用イベント生成処理の実行命令を送信する。管理用イベント生成処理とは、RAM64のキューに記録されている一連のサンプル値が示す音声波形に対し、ある周波数以下の周波数帯の成分を取り出すフィルタ処理と、その周波数よりさらに低いある周波数以下の周波数帯の成分を取り出すフィルタ処理を行い、それら2つのフィルタ処理により得られる値の比較判定を行うことにより、管理用イベントを生成する処理である。なお管理用イベントとは、後述する演奏イベントのタイミング調整処理において用いられる時間情報を生成するためのフラグである。以下、図6を参照しながら、管理用イベント生成処理を説明する。

【0068】

DSP63はCPU62より管理用イベント生成処理の実行命令を受信すると、RAM64のキューに記録されているサンプル値の末尾から一定数のサンプルを読み出す（ステップS11）。本実施形態においては、例として、DSP63はキューより44100組のサンプル値を読み出すこととする。以下、DSP63が管理用イベント生成処理においてキューから読み出す一連のサンプル値を「

未加工音声データ」と呼び、複数組の未加工音声データを区別する必要がある場合には、サンプル値（n）を末尾とする未加工音声データを「未加工音声データ（n）」のように表す。

【0069】

例えば、DSP63が管理用イベント生成処理の実行命令を受信した時点においてキューに記録されているサンプル値の末尾がサンプル値（50760）である場合、最初の未加工音声データは未加工音声データ（50760）、すなわちサンプル値（6601）～サンプル値（50760）である。

【0070】

続いて、DSP63は読み出した未加工音声データの各サンプル値の左右の値を相加平均することにより、ステレオのデータをモノラルのデータに変換する（ステップS12）。このモノラルへの変換処理は、このステップより後の処理のDSP63への負荷を軽減するための処理である。

【0071】

続いて、DSP63はモノラルに変換された一連のサンプル値について、絶対値を求める（ステップS13）。このステップにおける処理は、各サンプル値のパワーの代替値を求めるための処理である。絶対値はパワーを示す二乗値よりも値が小さく処理が容易であるため、本実施形態においては、各サンプル値の二乗値の代替値として絶対値が用いられる。従って、DSP63の処理能力が高い場合、このステップにおいて各サンプル値の絶対値の代わりに二乗値を算出してもよい。

【0072】

続いて、DSP63はステップS13において絶対値に変換された一連のサンプル値に対し、さらにローパスフィルタを用いたフィルタ処理を行う（ステップS14）。ステップS11～ステップS14の処理の結果得られる値を以下、「中期指標」と呼び、サンプル値（n）に対応する中期指標を「中期指標（n）」と呼ぶ。以下、例として、ステップS14において用いられるローパスフィルタの周波数は100Hzとする。中期指標（n）は、一連のサンプル値が示す音声波形のサンプル値（n）に対応する時点における中期的な変化の傾向を示す。す

なわち、一連のサンプル値が示す音声波形は短期的に上下に変化するが、ローパスフィルタによるフィルタ処理を施された一連のサンプル値は、先行する複数のサンプル値により値の変化が抑制される。その結果、一連の中期指標が示す音声波形からは短期的な変動要素が取り除かれ、中期的および長期的な変動要素のみが残る。ステップS14の処理により、中期指標(n)を末尾とする一連の中期指標、すなわち・・・中期指標(n-2)、中期指標(n-1)、中期指標(n)が作成される。DSP63はこれらの中間指標をRAM64に記録する(ステップS15)。

【0073】

続いて、DSP63はステップS14において得られる一連の中期指標に対し、樹形フィルタを用いたフィルタ処理を行う(ステップS16)。このステップにおける処理は、一連の中期指標が示す音声波形から、さらに、一定周波数以下の低周波数帯の成分を取り出すための処理である。このステップはステップS14で用いるローパスフィルタの周波数よりも低い周波数のローパスフィルタを用いる処理と同様であるが、樹形フィルタはローパスフィルタと比較して、通常、DSP63への負荷が小さいため、本実施形態においてはローパスフィルタが樹形フィルタで代用されている。

【0074】

図7は、ステップS16において利用可能な樹形フィルタの一例について、その構成を示している。図7において、四角形で示される処理は遅延処理を示し、 z^{-k} におけるkは、その遅延処理における遅延時間が(サンプリング周期×k)であることを意味する。前述のとおり、音楽CDのサンプリング周波数は44100Hzであるため、サンプリング周期は1/44100秒である。一方、三角形で示される処理は乗算処理を示し、三角形の中に示される値が乗算の係数を示す。以下、例として、k=22050とし、その結果、ステップS16のフィルタ処理により、1Hzより高い周波数帯の成分がおおよそ取り除かれる。すなわち、ステップS16の処理の結果得られる一連の値が示す音声波形は、一連の中期指標が示す音声波形から中期的な変動要素を取り除き、長期的な変動要素のみを取り出したものとなる。

【0075】

続いて、DSP63は、図6のステップS16のフィルタ処理により得られる一連の値のそれぞれに対し、正の定数hを乗ずる（ステップS17）。このhによる乗算処理は、次のステップS18における比較判定処理において肯定的な結果が得られる頻度を調整する処理であり、一般的にhの値が小さいと比較判定処理における肯定的な結果が得られる時間間隔が狭くなる。この時間間隔は、広すぎると以下のステップS21における管理用イベントの生成の時間間隔が広くなり、その結果、後述する演奏イベントのタイミング調整処理における精度が低くなる。一方、比較判定処理における肯定的な結果が得られる時間間隔が狭すぎると、以下のステップS20の処理により次々とそれらの肯定的な結果がキャンセルされてしまうため、やはり管理用イベントの生成の時間間隔が広くなり、その結果、演奏イベントのタイミング調整処理における精度が低くなる。従って、hの値としては、適度の頻度で管理用イベントが生成される値が経験的に用いられる。

【0076】

ステップS17の処理の結果得られる値を以下、「長期指標」と呼び、サンプル値（n）に対応する長期指標を「長期指標（n）」と呼ぶ。ステップS17の処理により、長期指標（n）を末尾とする一連の長期指標、すなわち・・・長期指標（n-2）、長期指標（n-1）、長期指標（n）が作成される。DSP63はこれらの長期指標をRAM64に記録する（ステップS18）。

【0077】

続いて、DSP63は、RAM64から、中期指標（n）と長期指標（n）を読み出し、中期指標（n）が長期指標（n）以上であるか否かの比較判定を行う（ステップS19）。この比較判定処理は、未加工音声データが示す音声波形が、サンプル値（n）に対応する時点において、中期的に大きく変動したことを示す。すなわち、楽曲の音声波形において、1Hz～100Hzの周波数帯に含まれる音のボリュームが急速に大きくなると、中期指標の値が長期指標の値を上回り、ステップS19における比較判定処理からは肯定的な結果（以下、「Yes」と呼ぶ）が得られる。

【0078】

ステップS19の比較判定処理によりY_esが得られると、DSP63は基準時点Pを基準として計時されるその時点の時間情報をRAM64に記録する。続いて、DSP63はRAM64から、過去に行われたステップS19における比較判定処理でY_esが得られた時間情報の記録を読み出し、過去の期間 τ にY_esが得られた記録があるか否かを判定する（ステップS20）。このステップS20における判定処理は、短い時間間隔でステップS19における比較判定処理の結果がY_esとなる場合、次のステップS21において、同じく短い時間間隔で管理用イベントを生成することを防ぐための処理である。もし短い時間間隔で管理用イベントが生成されると、後述する演奏イベントのタイミング調整処理において、記録されている管理用イベントと、新たに得られる音声データにより生成される管理用イベントとを正しく対応づけることが困難となる。ステップS20の処理により、期間 τ 以下の時間間隔で近接して管理用イベントが生成されることが回避される。 τ の値としては、生成される管理用イベントが適度の時間間隔となる値が経験的に用いられる。なお、ステップS20における最初の判定処理においては、先行するステップS19の比較判定処理がないため、判定結果はNoとなる。

【0079】

ステップS20における判定処理により否定的な結果（以下、「No」と呼ぶ）が得られた場合、DSP63は上記の一連の処理の結果、未加工音声データの示す音声波形が、サンプル値（n）に対応する時点で所定の条件を満たしたこと示す管理用イベントをCPU62に送信する（ステップS21）。

【0080】

ステップS19における比較判定の結果がNoであった場合、ステップS20における判定の結果がY_esであった場合、およびステップS21の処理を終了した場合、DSP63はCPU62がCDドライブ1から新たなサンプル値、すなわちサンプル値（n+1）を受信し、そのサンプル値がRAM64のキューに記録されるまで待機する。サンプル値（n+1）がキューに記録されると（ステップS22）、DSP63はサンプル値（n+1）を末尾とする未加工音声データ

タ ($n + 1$) に対し、上述したステップ S 1 1 以下の処理を行う。

【0081】

以上のステップ S 1 1～ステップ S 2 2 の処理は、音声データ N A の最後のサンプル値を末尾とする未加工音声データに対する処理が終了するまで続けられる。以上が管理用イベント生成処理の説明である。

【0082】

図 8 は、実際の音声データに対し、管理用イベント生成処理を行った場合の管理用イベント生成の様子を示す図である。なお、この図の作成においては、ステップ S 1 4 のローフィルタとして、1 段の I I R (Infinite Impulse Response) フィルタが用いられている。また、ステップ S 1 7 における定数 h としては 4 を、またステップ S 2 0 における期間 τ としては 0.55 秒が用いられている。

【0083】

図 8 において例示されている 2 つの管理用イベントのはじめの方に関し、管理用イベントが生成される時点 A のすぐ後に、中期指標が長期指標を上回る時点として B および C があるが、管理用イベントは A においてのみ生成される。なぜなら、B は A から所定の期間、すなわち 0.55 秒を経過しておらず、また C は B から所定の期間を経過していないため、ステップ S 2 0 の判定の結果が Yes となり、ステップ S 2 1 の処理が行われないためである。

【0084】

【1. 2. 1. 7 : 演奏イベントの生成】

上述した D S P 6 3 による管理用イベント生成処理が行われる一方で、ユーザはピアノ 3 1 を用いた演奏を開始する。すなわち、ユーザは発音部 4 から出力される楽曲 N に合わせて、ピアノ 3 1 の打鍵およびペダル操作を行う。ユーザのピアノ 3 1 を用いた演奏の情報は、キーセンサ 3 2 およびペダルセンサ 3 3 にて鍵およびペダルの動きとして検出され、M I D I イベント制御回路 3 4 にて演奏イベントに変換された後、コントローラ部 6 に送信される。

【0085】

【1. 2. 1. 8 : イベントデータの記録】

上述したように、CPU62は音声データNAの再生の間、DSP63から管理用イベントを、また自動演奏ピアノ部3のMIDIイベント制御回路34から演奏イベントをそれぞれ受信する。

【0086】

図9は、管理用イベントの生成と演奏イベントの生成の時間的な関係を示す図である。この場合、CPU62は基準時点Pから、1. 51秒、2. 38秒、4. 04秒・・・の時間が経過した時点で管理用イベントを受信している。また、CPU62は基準時点Pから、2. 11秒、2. 62秒、3. 60秒・・・の時間が経過した時点で演奏イベントを受信している。なお、例えば1. 78秒の時点で中期指標が長期指標を上回っているが、1. 51秒の時点で中期指標が長期指標を上回った後、0. 55秒を経過していないので、1. 78秒の時点で管理用イベントは受信されない。

【0087】

CPU62は管理用イベントを受信すると、管理用イベントを示すシステムエクスクルーシブイベントを生成し、生成したシステムエクスクルーシブイベントに対し、基準時点Pを基準とする管理用イベントの受信時刻を示す時間情報をデルタタイムとして付加し、RAM64に記録する。同様に、CPU62は演奏イベントを受信すると、受信した演奏イベントに対し、基準時点Pを基準とする演奏イベントの受信時刻を示す時間情報をデルタタイムとして付加し、RAM64に記録する。

【0088】

【1. 2. 1. 9：楽曲の末尾の特定】

CDドライブ1は、音声データNAの最後のサンプル値をコントローラ部6に送信すると、音楽CD-Aの再生を停止する。CPU62は、音声データNAの最後のサンプル値を受信し、そのサンプル値を末尾とする未加工音声データに対する管理用イベント生成処理を終了すると、続いて音声データNAにおける楽曲Nの実際の末尾を特定する処理を行う。

【0089】

CPU62が音声データNAの最後のサンプル値を末尾とする未加工音声データ

タに対する管理用イベント生成処理を終了した時点で、キューには音声データNAの最後のサンプル値を末尾とする1323000組のサンプル値が、各サンプル値の受信時刻を示す時間情報と共に記録されている。例えば、音声データNAの末尾のサンプル値がサンプル値(7673399)である場合、キューにはサンプル値(6350400)～サンプル値(7673399)が、それぞれのサンプル値に対応する時間情報と共に記録されている。

【0090】

CPU62はまず、キューに記録されている末尾のサンプル値を読み出し、読み出したサンプル値の左右いずれかの絶対値が予め定められた閾値である1000を超えているか否かの判定処理、すなわち閾値判定処理を行う。閾値判定処理において否定的な結果が得られた場合、CPU62はキューの末尾から2番目に記録されているサンプル値を読み出して、読み出したサンプル値に対し閾値判定処理を行う。CPU62は閾値判定処理により肯定的な結果が得られるまで、キューの末尾から先頭に向けて順次サンプル値を読み出し、同様の処理を繰り返す。

【0091】

例えば、サンプル値(7673399)を末尾とする音声データNAに関して、サンプル値(7634297)において初めて、R(7634297)もしくはL(7634297)の絶対値が1000を超えるとすると、サンプル値(7634298)～サンプル値(7673399)に対するCPU62の閾値判定処理は否定的な結果を得る。これは音声データNAの末尾に約0.89秒間の沈黙もしくは沈黙に近い部分が存在することを示す。

【0092】

上述した閾値判定処理により肯定的な結果を与えるサンプル値を以下、サンプル値(Z)とする。サンプル値(Z)は楽曲Nの実際の末尾に対応するサンプル値である。CPU62は閾値判定処理の肯定的な結果を得ることによりサンプル値(Z)を特定すると、それ以降の閾値判定処理は行わない。

【0093】

[1. 2. 1. 10 : 末尾参照用加工済音声データの生成]

続いて、CPU62は以下に説明する末尾参照用加工済音声データの生成処理を行う。末尾参照用加工済音声データの生成処理とは、キューに記録されている音声データに対し、上述した相関判定用データ生成処理を行うことにより、後述する相関判定処理を行うための音声データを生成する処理である。

【0094】

以下、図10を参照しつつ、末尾参照用加工済音声データ生成処理を説明する。なお、以下の説明においてはキューに記録されているサンプル値の先頭をサンプル値(W)とする。また、説明をより具体的にするため、例としてW=6350400、Z=7634297とする。これは、キューにはサンプル値(6350400)～サンプル値(7673399)が記録されており、また楽曲Nの実際の末尾に対応するサンプル値がサンプル値(7634297)であることを意味する。また、サンプル値(n)を末尾とする65536組の一連のサンプル値を「参照用未加工音声データ(n)」と呼ぶ。

【0095】

CPU62はまず、カウンタiおよびカウンタjを生成し、初期値としてi=Z=7634297、j=0をセットする(ステップS31)。続いて、CPU62はDSP63に対し、参照用未加工音声データ(i-j)に対する相関判定用データ生成処理の実行命令を送信する。相関判定用データ生成処理は図4を用いて既に説明した処理であるので、説明を省略する。この相関判定用データ生成処理の実行命令に応じ、DSP63は参照用未加工音声データ(i-j)を加工して得られる256の一連のサンプル値をRAM64に記録する(ステップS32)。以下、この256のサンプル値を「参照用加工済音声データ」と呼び、参照用未加工音声データ(n)から得られる参照用加工済音声データを「参照用加工済音声データ(n)」と呼ぶこととする。なお、最初にステップS32が実行される際には、i-j=7634297であるので、参照用加工済音声データ(7634297)がRAM64に記録される。

【0096】

続いて、CPU62はj=881999であるか否かを判定する(ステップS33)。この判定処理は、音声データNAの末尾の20秒間分に対応する音声デ

ータに関し、末尾を1サンプルずつずらしながら一連の65536組のサンプル値を読み出すことにより、882000組の参照用未加工音声データを順に読み出すために行われる処理である。

【0097】

ここで、 $j = 0$ であるので、ステップS33における判定処理の結果はNoとなる。その場合、CPU62は j の値を1増加させ（ステップS34）、上述したステップS32の処理に戻る。ステップS32、ステップS33、ステップS34の一連の処理は881999回繰り返され、ステップS32の処理は882000回実行される。その結果、RAM64には参照用加工済音声データ（7634297）、参照用加工済音声データ（7634296）、・・・、参照用加工済音声データ（6752298）が記録される。

【0098】

882000回目のステップS33において、 $j = 881999$ であるため、判定処理の結果がYesとなる。続いて、CPU62はDSP63に対し、RAM64に記録されている参照用加工済音声データ（i）の参照用加工済音声データ（i-j）に対する相関判定処理の実行命令を送信する（ステップS35）。相関判定処理とは、2組の音声データの類似性を判定する処理である。以下、図11を参照しつつ、相関判定処理を説明する。

【0099】

CPU62は、DSP63に相関判定処理の実行命令を送信する際、類似性の判定における参照元となる音声データ（以下、「参照元音声データ」と呼ぶ）と、参照先となる音声データ（以下、「参照先音声データ」と呼ぶ）を指定する。この場合、参照元音声データは参照用加工済音声データ（i）であり、参照先音声データは参照用加工済音声データ（i-j）である。以下、参照元音声データに含まれる256のデータを $X(0) \sim X(255)$ と表す。一方、参照先音声データに含まれる256のデータを $Y_m(0) \sim Y_m(255)$ と表す。ただし、 $m = i - j$ であり、参照先音声データが生成された際に用いられた参照用未加工音声データの末尾に対応するサンプル値の番号を示す。

【0100】

DSP63はCPU62から相関判定処理の実行命令を受信すると、RAM64から参照用加工済音声データ(*i*)を参照元音声データとして、また参照用加工済音声データ(*i-j*)を参照先音声データとして読み出し、以下の式(2)および式(3)で表される判定処理を行う(ステップS51)。なお、最初に実行されるステップS51の処理においては、*i*=7634297、*m*=*i-j*=6752298である。

【数2】

$$\frac{\sum_{i=0}^{255} (X(i) \times Y_m(i))}{\sum_{i=0}^{255} (X(i)^2)} \geq p \quad \dots \quad (2)$$

【数3】

$$\frac{\left\{ \sum_{i=0}^{255} (X(i) \times Y_m(i)) \right\}^2}{\sum_{i=0}^{255} (X(i)^2) \times \sum_{i=0}^{255} (Y_m(i)^2)} \geq q \quad \dots \quad (3)$$

【0101】

式(2)の左辺はX(0)～X(255)とY_m(0)～Y_m(255)を順に並べて同じ番号のデータを対にした時、各対のデータの値が合致すればする程、大きな値をとり、その値は1に近づく。以下の説明において、この左辺の値を「絶対相関指標」と呼ぶ。なお、pの値は0～1の範囲で任意に変更が可能で、異なる版における同じ楽曲の同じ部分に対応する音声データに対し相関判定用データ生成処理を行うことにより得られる2組のデータに関しては上記の式(2)による判定の結果がY_{es}となり、たとえ類似していても楽曲における異なる部分に対応する音声データから得られる2組のデータに関しては、式(2)による判定の結果がNoとなるよう、経験的に定められる。

【0102】

式(3)の左辺は0～1の範囲で値をとり、X(0)～X(255)が示す音声波形とY_m(0)～Y_m(255)が示す音声波形の形状が相似形に近いほど、

1に近づく。以下の説明において、この左辺の値を「相対相関指標」と呼ぶ。上記の絶対相関指標は、 $X(0) \sim X(255)$ および $Y_m(0) \sim Y_m(255)$ が示す音声波形の形状が類似していても、 $X(0) \sim X(255)$ および $Y_m(0) \sim Y_m(255)$ が示す音声波形が示す音声波形のレベルが互いに異なると、値が変動する。それに対し、相対相関指標は $X(0) \sim X(255)$ および $Y_m(0) \sim Y_m(255)$ が示す音声波形のレベルの大小に影響を受けず、形状が類似していれば1に近い値をとるため、音楽CDの異なる版において録音レベルが違う場合であっても、式(3)の判定は Yes を与える。qの値は0~1の範囲で任意に変更が可能で、pと同様に経験的に定められる。

【0103】

ステップS51における2つの判定処理の結果の両方がYesである場合、DSP63は以下の式(4)および式(5)で表される判定処理を行う(ステップS52)。

【数4】

$$\frac{d \sum_{i=0}^{255} (X(i) \times Y_m(i))}{dm} = 0 \quad \dots (4)$$

【数5】

$$\frac{d^2 \sum_{i=0}^{255} (X(i) \times Y_m(i))}{d^2 m} < 0 \quad \dots (5)$$

【0104】

式(4)の左辺は $X(0) \sim X(255)$ と $Y_m(0) \sim Y_m(255)$ の積和のサンプル値(m)における変化率である。以下の説明において、 $X(0) \sim X(255)$ と $Y_m(0) \sim Y_m(255)$ の積和を「相関値」と呼ぶ。相関値は、 $X(0) \sim X(255)$ と $Y_m(0) \sim Y_m(255)$ を順に並べて同じ順番のデータを対にした時、各対のデータの値が近似すればする程、大きな値となる。そ

して、相関値が極値をとる場合、 $X(0) \sim X(255)$ と $Y_m(0) \sim Y_m(255)$ に関する相関値を m について時系列的に並べた際の変化率が 0 となる。従って、式(4)による判定処理は、相関値が極値であるかどうかを判定する処理である。また、式(5)による判定処理は、その極値が極大値であることを判定する処理である。

【0105】

より正確に説明すると、本実施形態においては $X(0) \sim X(255)$ および $Y_m(0) \sim Y_m(255)$ は離散値であるため、式(4)の左辺が厳密に 0 を取ることは極めてまれである。従って、ステップ S 5 2 の判定処理は、実際には次のように行われる。まず、DSP 6 3 は $X(0) \sim X(255)$ と $Y_m(0) \sim Y_m(255)$ の積和と、 $X(0) \sim X(255)$ と $Y_{m-1}(0) \sim Y_{m-1}(255)$ の積和との差を取る。その値を以下、 D_m と呼ぶ。続いて、DSP 6 3 は D_{m-1} が 0 より大きく、かつ D_m が 0 以下であるか否かを判定する。ここで D_{m-1} が 0 より大きく、かつ D_m が 0 以下である場合、相関値の変化率は D_m において正の値から 0 になるか、0 をまたがって変化しているので、その時点における相関値は極大値もしくは極大値の近似値である。従って、その場合、ステップ S 5 2 の判定処理の結果は Yes となる。それ以外の場合、ステップ S 5 2 の判定処理の結果は No となる。

【0106】

ステップ S 5 2 の判定処理の結果が Yes である場合、DSP 6 3 は CPU 6 2 に対し、参照先音声データが参照元音声データに極めて類似していることを示す成功通知を送信する（ステップ S 5 3）。

【0107】

一方、ステップ S 5 1 の判定処理の結果が No である場合、もしくはステップ S 5 2 の判定処理の結果が No である場合、DSP 6 3 は CPU 6 2 に対し、参照先音声データが参照元音声データにあまり類似していないことを示す失敗通知を送信する（ステップ S 5 4）。

【0108】

図 1 2 は、実際の音声データのサンプルに対し、ステップ S 5 1 およびステッ

PS52の判定処理において用いられる計算式の値を算出し、グラフに示したものである。なお、このグラフの作成においては、図4におけるステップS3においては周波数25Hzのハイパスフィルタとして、1段のIIRフィルタを、またステップS5における擡形フィルタの定数としては、 $k = 4410$ および $f = 1$ を、また、ステップS6においては周波数25Hzのローパスフィルタとして、1段のIIRフィルタを用いている。さらに、図11のステップS51における判定式の定数としては、 $p = 0.5$ 、 $q = 0.8$ を用いている。

【0109】

図12の上段のグラフには、 m に対する、式(2)の左辺の分子の値と、左辺の分母を右辺に移項した式の値が示されている。図12の中段のグラフには、 m に対する、式(3)の左辺の分子の値と、左辺の分母を右辺に移項した式の値が示されている。また、図12の下段のグラフには、式(4)の左辺の値が示されている。

【0110】

図12によれば、 m の値が区間Aに位置する場合において、式(2)の左辺の分子の値が、左辺の分母を右辺に移項した式の値と等しいか、それを超えるため、式(2)の条件が満たされる。さらに、区間Aのなかで、 m の値が区間Bに位置する場合において、式(3)の左辺の分子の値が、左辺の分母を右辺に移項した式の値と等しいか、それを超えるため、式(3)の条件が満たされる。その結果、ステップS51の判定処理において、Yesの結果が得られる。区間Bのなかで、 m の値が矢印Cで示される値をとる場合、式(4)の左辺の値が正の値から0となり、式(5)の条件も満たされるので、ステップS52の判定処理において、Yesの結果が得られる。

以上が相関判定処理の説明である。

【0111】

図10のステップS35において、上述したようにDSP63は相関判定処理の結果として、成功通知もしくは失敗通知をCPU62に送信する。CPU62は、DSP63からいずれかの通知を受信すると、その通知が成功通知であるか失敗通知であるかを判定する(ステップS36)。

【0112】

通常、相関判定処理は楽曲の同じ部分に対応する2組の音声データを用いて行われる場合にのみ、式(2)～式(5)で示した条件が全て満たされ、その結果、CPU62は成功通知を受信する。ステップS35が最初に実行される場合、相関判定処理に用いられる参照元音声データは参照用加工済音声データ(7634297)であり、参照先音声データは参照用加工済音声データ(6752298)であるため、CPU62は通常、ステップS35の結果として、失敗通知を受信する。その場合、CPU62はjの値を1減少させ(ステップS37)、上述したステップS35の処理に戻る。

【0113】

音声データNAの末尾の20秒間に関し、参照元音声データに対応する音声波形と極めて類似した音声波形が他に存在しない場合、ステップS35、ステップS36、ステップS37の一連の処理は881999回繰り返され、ステップS35の処理は882000回実行される。その間、iの値は一定であるため、参照元音声データは変化しないが、参照先音声データはjの値が減少するため、変化する。例えば、参照元音声データが参照用加工済音声データ(7634297)である場合、参照先音声データは参照用加工済音声データ(6752298)、参照用加工済音声データ(6752299)、参照用加工済音声データ(6752300)、・・・と変化する。882000回目のステップS35の処理においては、j=0となるため、参照先音声データは参照元音声データと一致し、その結果、CPU62はステップS35の結果として成功通知を受信する。

【0114】

CPU62がステップS35の結果として成功通知を受信すると、続くステップS36の判定処理の結果はYesとなる。その場合、CPU62はjの値が0であるか否かを判定する(ステップS38)。上述したように、もし音声データNAの末尾の20秒間に関し、参照元音声データに対応する音声波形と極めて類似した音声波形が他に存在しないならば、j=0となる。その場合、CPU62はその際の参照元音声データである参照用加工済音声データ(i)をRAM64に記録する。さらに、CPU62は、RAM64のキューにおいてサンプル値(

i) に対応付けられて記録されている時間情報 (i) を読み出し、RAM64に記録する (ステップS39)。以下、ステップS39の処理においてRAM64に記録される参照用加工済音声データを「末尾参照用加工済音声データ」、また時間情報 (i) を「末尾時間情報」と呼ぶ。例えば、初めて実行されるステップS38における判定処理の結果がYesとなつた場合、 $i = 7634297$ であるので、末尾時間情報はサンプル値 (7634297) に対応する時間情報 (7634297)、すなわち約173.11秒となる。

【0115】

一方、音声データNAの末尾の20秒間に関し、参照元音声データに対応する音声波形と極めて類似した音声波形が他に存在する場合がある。その場合、上述したステップS35、ステップS36、ステップS37の一連の処理が881999回繰り返される前に、ステップS35において式(2)～式(5)で示した条件が全て満たされ、その結果、CPU62は成功通知を受信する。その結果、ステップS36における判定結果がYesとなり、ステップS38に処理が移るが、その場合、まだ $j = 0$ となつてないため、ステップS38の判定結果はNoとなる。

【0116】

ステップS38の判定結果がNoである場合、CPU62は参照元音声データである参照用加工済音声データ (i) を変更し、ステップS32～ステップS38の処理を繰り返す。そのために、まずCPU62は、 $i = W + 65536 + 881999 = W + 947535 = 7297935$ であるか否かを判定する (ステップS40)。もし、参照元音声データ (i) を生成する際に用いられる参照用未加工音声データ (i) が、キューの先頭から数えて882000番目の参照用未加工音声データにあたると、ステップS40の判定結果はYesとなる。ステップS40が最初に実行される際、 $i = Z = 7634297$ であるため、ステップS40の判定結果はNoとなる。

【0117】

ステップS40の判定結果がNoである場合、CPU62は i を1減少させる。また、CPU62は j に881999を設定する (ステップS41)。その後

、CPU62は処理をステップS32に戻す。ステップS32においては、iが1減少され、さらにj = 881999であるため、参照用未加工音声データ(i - j)として、先行するステップS32において既に相関判定用データ生成処理が行われた参照用未加工音声データより、キューにおいて1サンプル値だけ先頭側に位置する参照用未加工音声データがキューより読み出される。例えば、ステップS41の処理が最初に実行された直後に実行されるステップS32においては、i = 7634296であるため、参照用未加工音声データ(6752297)に対し、相関判定用データ生成処理が行われる。その結果、既にRAM64に記録されている参照用加工済音声データ(7634297)～参照用加工済音声データ(6752298)に加え、新たに参照用加工済音声データ(6752297)がRAM64に記録される。

【0118】

続くステップS33の判定処理において、j = 881999であるため、判定結果はYesとなる。続いて、CPU62はステップS35、ステップS36、ステップS37の一連の処理を繰り返す。ステップS35の相関判定処理の結果として成功通知がCPU62により受信され、ステップS36の判定処理の結果がYesとなると、CPU62はステップS38の判定処理を行う。

【0119】

既に上述したように、サンプル値(i)を末尾とする音声データに対応する音声データNAの20秒間の音声波形に関し、参照元音声データに対応する音声波形と極めて類似した音声波形が他に存在しない場合、j = 0となる。従って、CPU62はステップS39を実行し、その結果、RAM64には末尾参照用加工済音声データおよび末尾時間情報が記録される。一方、サンプル値(i)を末尾とする音声データに対応する音声データNAの20秒間の音声波形に関し、参照元音声データに対応する音声波形と極めて類似した音声波形が他に存在しない場合、j > 0である。従って、CPU62はステップS40の判定処理を行う。

【0120】

ステップS38の判定処理の結果が繰り返しNoとなると、ステップS40およびステップS41の処理が繰り返されることにより、iの値は減少する。例え

ば、キューに記録されているサンプル値が示す音声波形が一定の値をとるような場合、ステップS38の判定結果はY e sになることはない。その結果、 $i = W + 947535 = 7297935$ になり、ステップS40の判定結果がY e sとなる。これは、キューに記録されているサンプル値により得られる音声データNAの末尾付近の音声波形において、特徴的な形状を持つ音声波形が存在せず、その結果、末尾参照用加工済音声データの取得が失敗したことを意味する。従って、CPU62は操作表示部5にエラーメッセージを表示させ（ステップS42）、一連の処理を終了する。

以上が、末尾参照用加工済音声データの生成処理である。

【0121】

【1. 2. 1. 11 : SMFのFDへの記録】

上記の末尾参照用加工済音声データの生成処理におけるステップS39の処理により、RAM64に末尾参照用加工済音声データおよび末尾時間情報が記録された場合、CPU62は続いてSMFのFDへの記録処理を行う。

【0122】

まずCPU62は、RAM64に記録されている以下のデータを読み出す。

- (1) 先頭参照用加工済音声データ
- (2) 先頭時間情報
- (3) イベントデータ
- (4) 末尾参照用加工済音声データ
- (5) 末尾時間情報

【0123】

続いて、CPU62は読み出したこれらのデータを用いて、SMFのトラックチャックを生成する。さらに、CPU62は作成したトラックチャックに対し、これに応じたヘッダチャックを付加し、SMFを生成する。

【0124】

図13は、CPU62が生成するSMFの概略を示す図である。トラックチャックのデータ部分の先頭には、まず、先頭参照用加工済音声データと先頭時間情報とを含むシステムエクスクルーシブイベントおよび、末尾参照用加工済音声データ

タと末尾時間情報を含むシステムエクスクルーシブイベントが、対応するデルタタイムと共に記録されている。これらのシステムエクスクルーシブイベントのデルタタイムは任意であるが、本実施形態においては0.00秒とする。上述したシステムエクスクルーシブイベントに続き、管理用イベントおよび演奏イベントが、対応するデルタタイムと共に、デルタタイムの順に記録されている。

【0125】

CPU62は、SMFの生成を完了すると、生成したSMFを書込命令と共にFDドライブ2に送信する。FDドライブ2はCPU62から書込命令およびSMFを受信すると、セットされているFDにSMFを書き込む。

【0126】

図14は音声データNAと、SMFに書き込まれる時間情報およびデルタタイムを整理した図である。先頭参照用加工済音声データと共にSMFに記録される先頭時間情報は、音声データNAのうち、楽曲Nが実際に開始する時刻を示している。また、末尾参照用加工済音声データと共にSMFに記録される末尾時間情報は、音声データNAのうち、楽曲Nが実際に終了する時刻を示している。

【0127】

[1. 2. 2: 再生動作]

続いて、上述した方法によって記録されたSMFを用いて、音楽CDに含まれる音声データとSMFのMIDIデータを同期再生させる際の動作を説明する。以下の再生動作において用いられる音楽CDは、上述した記録動作において用いられた楽曲Nの音声データを含んでいるが、版が異なっており、音楽データの再生開始の時点から実際に楽曲Nが開始するまでの時間や、音声データの示す音声波形のレベルが異なっている。また、楽曲Nのマスターデータから音楽CD用のデータが作成される際、音声データに対し音響効果等に関する編集が加えられているため、この音楽CDに含まれる楽曲Nの音声データは、音楽CD-Aに含まれる音声データNAと、その内容にわずかの差異がある。以下に説明する再生動作において用いられる音楽CDを音楽CD-Bと呼び、音楽CD-Bに含まれる楽曲Nの音声データを音声データNBと呼ぶ。

【0128】

[1. 2. 2. 1 : 再生の開始操作]

ユーザは、音楽CD-BをCDドライブ1に、またSMFの記録されたFDをFDドライブ2にセットする。続いて、ユーザは音声データNBとSMFの同期再生開始に対応する操作表示部5のキーパッドを押下する。操作表示部5は押下されたキーパッドに対応する信号をコントローラ部6に出力する。

【0129】

CPU62は操作表示部5から同期再生開始を指示する信号を受信すると、まずFDドライブ2に対しSMFの送信命令を送信する。FDドライブ2はこのSMFの送信命令に応じて、FDからSMFを読み出し、読み出したSMFをコントローラ部6に送信する。CPU62はFDドライブ2からSMFを受信し、受信したSMFをRAM64に記録する。

【0130】

[1. 2. 2. 2 : 演奏イベントのタイミング調整]

CPU62はSMFをRAM64に記録すると、演奏イベントのタイミング調整処理を行う。演奏イベントのタイミング調整処理とは、SMFの記録時に用いられた音声データNAと、SMFの再生時に用いられる音声データNBとの間の楽曲開始前の沈黙時間および楽曲終了後の沈黙時間の差、すなわち楽曲Nの開始時点および楽曲Nの再生速度の差により生ずる、音声データNBの楽曲Nの再生とSMFのMIDIデータの再生とのタイミングのずれを調整する処理である。以下、図15を用いて、演奏イベントのタイミング調整処理を説明する。

【0131】

CPU62は、カウンタiを生成し、iに65535をセットする（ステップS61）。続いて、CPU62はCDドライブ1に音声データNBの再生命令を送信し、CDドライブ1は音声データNBのサンプル値を、1/44100秒ごとに先頭から順にコントローラ部6に送信する。以下の説明において、音声データNBのサンプル値を、音声データNAの場合と同様に、先頭から、サンプル値(0)、サンプル値(1)、・・・のように表す。

【0132】

CPU62は、CDドライブ1から音声データNBの最初のサンプル値、すな

わちサンプル値 (0) を受信すると、クロックから得られるクロック信号に基づき、その時点（以下、「基準時点Q」と呼ぶ）を基準とする計時を開始する。

【0133】

CPU62は受信したサンプル値を、受信時刻を示す時間情報と共に、順次、RAM64のキューに記録する。この時間情報は、基準時点Qからの経過時間である。以下、上述した記録動作の場合と同様に、サンプル値 (n) に対応する時間情報を時間情報 (n) と呼ぶ。また、キューに記録可能なサンプル値の数は1323000組とする。

【0134】

CPU62は、サンプル値 (i) をキューに記録すると、キューに記録されているサンプル値 (i) を末尾とする65536組のサンプル値に対する相関判定用データ生成処理の実行命令を、DSP63に送信する。以下、サンプル値 (n) を末尾とする65536組の一連のサンプル値を、上述した記録動作の場合と同様に、参照用未加工音声データ (n) と呼ぶ。DSP63は、参照用未加工音声データ (i) に対する相関判定用データ生成処理を実行する（ステップS62）。ステップS62における相関判定用データ生成処理は、図4を用いて既に説明した処理と同様であるので、説明を省略する。なお、ステップS62が最初に実行される際、 $i = 65535$ であるので、参照用未加工音声データ (65535) に対し、相関判定用データ生成処理が行われる。ステップS62の処理の結果、RAM64には参照用加工済音声データ (i) が記録される。

【0135】

続いて、CPU62はDSP63に対し、先頭参照用加工済音声データの、参照用加工済音声データ (i) に対する相関判定処理の実行命令を送信する。DSP63は、RAM64に記録されたSMFに含まれる先頭参照用加工済音声データと、ステップS62においてRAM64に記録された参照用加工済音声データ (i) を読み出し、相関判定処理を実行する（ステップS63）。ステップS63における相関判定処理は、図11を用いて既に説明した処理と同様であるので、説明を省略する。ステップS63の処理の結果として、DSP63はCPU62に対し、失敗通知もしくは成功通知を送信する。

【0136】

CPU62はDSP63から失敗通知もしくは成功通知を受信すると、受信した通知が成功通知であるか否かを判定する（ステップS64）。音声データNBの先頭から楽曲Nが実際に開始することはまれであるので、通常、ステップS64が初めて実行される場合、判定処理の結果はN○となる。その場合、CPU62は $i = 65535 + 882000 = 947535$ であるか否かを判定する（ステップS65）。ステップS65における判定処理は、音声データNBの冒頭の20秒間分から得られる参照用加工済音声データ（i）の全てに関し、ステップS63における相関判定処理の結果が失敗であった場合、ステップS62～ステップS66の一連の処理の繰り返しを中止するための判定処理である。ステップS65が初めて実行される場合、 $i = 65535$ であるため、判定処理の結果はN○となる。CPU62はiを1増加させ（ステップS66）、ステップS62に処理を移動する。

【0137】

続いて、CPU62はステップS62～ステップS66の一連の処理を繰り返し、参照元音声データとしての先頭参照用加工済音声データに対し、参照先音声データを参照用加工済音声データ（65535）、参照用加工済音声データ（65536）、参照用加工済音声データ（65537）、・・・と順次、変更しながら、相関判定処理が実行される。その結果、通常は何回目かのステップS62において用いられる参照用未加工音声データ（i）が、先頭参照用加工済音声データに対応する楽曲Nの部分を示す音声データに達する。

【0138】

例えば、楽曲Nの先頭参照用加工済音声データに対応する部分が、音声データNBにおいてサンプル値（28740）を先頭とする一連の音声データであったとすると、ステップS62～ステップS66の一連の処理が28740回繰り返された後、28741回目のステップS62において相関判定用データ生成処理の対象となる参照用未加工音声データ（94275）は、音声データNBにおいて、楽曲Nの先頭参照用加工済音声データに対応する部分である。

【0139】

上記のように、ステップS62で用いられる参照用未加工音声データ(i)が、先頭参照用加工済音声データに対応する楽曲Nの部分を示す場合、ステップS62で生成される参照用加工済音声データ(i)と先頭参照用加工済音声データは、楽曲Nの同じ部分を示す音声データに対し、同じ相関判定用データ生成処理を施すことにより生成された音声データとなる。従って、これらの音声データの類似性は極めて高くなり、ステップS63の相関判定処理の結果として、CPU62は成功通知を受信する。その結果、ステップS64の判定処理の結果はYesとなる。

【0140】

ステップS64の判定結果がYesである場合、CPU62はCDドライブ1に、音声データNBの再生の停止命令を送信し、CDドライブ1は音声データNBの送信を停止する。続いて、CPU62は、 $(i - 65535) / 44100$ の値を求めるにより、参照用加工済音声データ(i)の先頭に対応する時間情報を算出する。例えば、ステップS64の判定処理の結果がYesとなった際、 $i = 94275$ であるとすると、 $(94275 - 65535) / 44100 =$ 約0.65であるから、参照用加工済音声データ(i)の先頭に対応する時間情報は約0.65秒となる。続いて、CPU62はRAM64に記録されているSMFから先頭時間情報を読み出し、参照用加工済音声データ(i)の先頭に対応する時間情報と、読み出した先頭時間情報を用いて、時間差を算出する。以下、この時間差を「トップオフセット」と呼ぶ。トップオフセットは、音声データNAの楽曲Nに対し、音声データNBの楽曲Nの開始時点が早い場合に負、遅い場合に正の値をとる。例えば、参照用加工済音声データ(i)の先頭に対応する時間情報が0.65秒であり、先頭時間情報が1.15秒である場合、トップオフセットは、 $0.65 - 1.15 = -0.50$ より、-0.50秒となる。CPU62は、トップオフセットをRAM64に記録する(ステップS67)。

【0141】

CPU62は、ステップS67においてトップオフセットの記録を行った後、RAM64に記録されたSMFから先頭時間情報および末尾時間情報を読み出し、その時間差を算出する。続いて、CPU62は、この時間差が示す時間に44

100を乗算することにより、時間差に相当するサンプル値の数を算出する。このサンプル値の数は、先頭参照用加工済音声データの先頭から、末尾参照用加工済音声データの末尾までの時間に対応するサンプル値の数である。そこで、CPU62は、このサンプル値の数から65536を減算することにより、先頭参照用加工済音声データの末尾から、末尾参照用加工済音声データの末尾までの時間に対応するサンプル値の数を求める。このようにして求められるサンプル値の数を以下、「V」とする。

【0142】

例えば、先頭時間情報が1.15秒、末尾時間情報が173.11秒である場合、先頭参照用加工済音声データの先頭から、末尾参照用加工済音声データの末尾までの時間は $173.11 - 1.15 = 171.96$ （秒）となる。この時間に対応するサンプル値の数は、 $171.96 \times 44100 = 7583436$ となる。従って、先頭参照用加工済音声データの末尾から、末尾参照用加工済音声データの末尾までの時間に対応するサンプル値の数、すなわちVは、 $7583436 - 65536 = 7517900$ となる。

【0143】

続いて、CPU62はカウンタjを生成し、 $j = i + V - 441000$ をセットする（ステップS68）。音声データNBにおいて、参照用未加工音声データ（i）は楽曲Nの先頭部分の約1.49秒に対応する音声データである。従って、楽曲Nの末尾部分の約1.49秒に対応する音声データは、参照用未加工音声データ（i+V）の前後に位置することが推定される。そして、参照用未加工音声データ（j）、すなわち参照用未加工音声データ（i+V-441000）は、参照用未加工音声データ（i+V）から10秒間だけ前に位置するデータである。

【0144】

続いて、CPU62はCDドライブ1に対し、音声データNBのサンプル値（ $j - 65535$ ）以降の再生命令を送信し、CDドライブ1はサンプル値（ $j - 65535$ ）、サンプル値（ $j - 65534$ ）、・・・を順次、コントローラ部6に送信する。CPU62は受信したサンプル値を、受信時刻を示す時間情報と

共に、順次、RAM64のキューに記録する。

【0145】

CPU62は、サンプル値(j)をキューに記録すると、DSP63に対し、参照用未加工音声データ(j)に対する相関判定用データ生成処理の実行命令を送信する。DSP63は、参照用未加工音声データ(j)に対する相関判定用データ生成処理を実行する(ステップS69)。ステップS69における相関判定用データ生成処理は、図4を用いて既に説明した処理と同様であるので、説明を省略する。ステップS69の処理の結果、RAM64には参照用加工済音声データ(j)が記録される。

【0146】

続いて、CPU62はDSP63に対し、末尾参照用加工済音声データの、参照用加工済音声データ(j)に対する相関判定処理の実行命令を送信する。DSP63は、RAM64に記録されたSMFに含まれる末尾参照用加工済音声データと、ステップS69においてRAM64に記録された参照用加工済音声データ(j)を読み出し、相関判定処理を実行する(ステップS70)。ステップS70における相関判定処理は、図11を用いて既に説明した処理と同様であるので、説明を省略する。ステップS70の処理の結果として、DSP63はCPU62に対し、失敗通知もしくは成功通知を送信する。

【0147】

CPU62はDSP63から失敗通知もしくは成功通知を受信すると、受信した通知が成功通知であるか否かを判定する(ステップS71)。音声データNBにおける楽曲Nの先頭から末尾までの時間が、音声データNAにおける楽曲Nの先頭から末尾までの時間と10秒間ずれることはまれであるので、通常、ステップS71が初めて実行される場合、判定処理の結果はN○となる。その場合、CPU62はjを1増加させ(ステップS72)、DSP63に対し、次の参照用未加工音声データに対する相関判定用データ生成処理の実行命令を送信する。ここで、jの値が音声データNBのサンプル値の総数よりも大きい場合、すなわち、参照用未加工音声データ(j-1)が既に音声データNBの最後に到達している場合、DSP63は参照用未加工音声データ(j)のキューからの読み出しに

失敗し、CPU62にエラー通知を送信する（ステップS73）。ステップS72が初めて実行される場合、サンプル値（j）はまだ音声データNBの最後のサンプル値に達していないので、CPU62はステップS73においてDSP63からエラー通知を受信することなく、DSP63は新たな参照用未加工音声データ（j）に対し相関判定用データ生成処理を実行する（ステップS69）。

【0148】

その後、CPU62はステップS69～ステップS73の一連の処理を繰り返し、参照元音声データとしての末尾参照用加工済音声データに対し、参照先音声データを参照用加工済音声データ（j）を順次変更しながら、相関判定処理が実行される。その結果、通常は何回目かのステップS69において用いられる参照用未加工音声データ（j）が、末尾参照用加工済音声データに対応する楽曲Nの部分を示す音声データに達する。その結果、ステップS69で生成される参照用加工済音声データ（j）と末尾参照用加工済音声データは、楽曲Nの同じ部分を示す音声データに対し、同じ相関判定用データ生成処理を施すことにより生成された音声データとなる。従って、これらの音声データの類似性は極めて高くなり、ステップS70の相関判定処理の結果として、CPU62は成功通知を受信する。その結果、ステップS71の判定処理の結果はYesとなる。

【0149】

ステップS71の判定結果がYesである場合、CPU62はCDドライブ1に、音声データNBの再生の停止命令を送信し、CDドライブ1は音声データNBの送信を停止する。続いて、CPU62は、 $j / 44100$ の値を求めるにより、参照用加工済音声データ（j）の末尾に対応する時間情報を算出する。例えば、ステップS71の判定処理の結果がYesとなった際、 $i = 7651790$ であるとすると、 $7651790 / 44100 =$ 約173.51であるから、参照用加工済音声データ（j）の末尾に対応する時間情報は約173.51秒となる。続いて、CPU62はRAM64に記録されているSMFから末尾時間情報を読み出し、参照用加工済音声データ（j）の末尾に対応する時間情報と、読み出した末尾時間情報を用いて、時間差を算出する。以下、この時間差を「エンドオフセット」と呼ぶ。エンドオフセットは、音声データNAの楽曲Nに対

し、音声データN Bの楽曲Nの終了時点が早い場合に負、遅い場合に正の値をとる。例えば、参照用加工済音声データ(j)の末尾に対応する時間情報が173.51秒であり、末尾時間情報が173.11秒である場合、トップオフセットは、173.51 - 173.11 = 0.40より、0.40秒となる。CPU62は、エンドオフセットをRAM64に記録する(ステップS74)。

【0150】

CPU62は、ステップS74におけるエンドオフセットの記録を終了すると、RAM64に記録されたSMFに、トップオフセットおよびエンドオフセットを含むシステムエクスクルーシブイベントを追加する(ステップS75)。図16は、トップオフセットおよびエンドオフセットが追加された後のSMFの内容を示した図である。トップオフセットおよびエンドオフセットを含むシステムエクスクルーシブイベントのデルタタイムは任意であるが、本実施形態においては0.00秒とする。

【0151】

続いて、CPU62は、SMFから演奏イベントに関するデータを全て読み出し、それらのデルタタイムを用いた以下の式(6)に従い、調整後のデルタタイムを算出し、その結果をRAM64に記録する(ステップS76)。ただし、式(6)において、d:調整後のデルタタイム、D:調整前のデルタタイム、N_T:先頭時間情報、N_E:末尾時間情報、O_T:トップオフセット、O_E:エンドオフセット、である。

【数6】

$$d = (N_T + O_T) + (D - N_T) \times \frac{(N_E + O_E) - (N_T + O_T)}{(N_E - N_T)} \dots (6)$$

【0152】

式(6)において、第1項の(N_T+O_T)は音声データN Bにおいて、音声データN Bの最初のサンプル値の再生タイミングを基準とした楽曲Nの開始時点を示す。(D-N_T)は、音声データN Aにおいて、楽曲Nの開始時点を基準とした演奏イベントの実行タイミングを示す。(N_E+O_E) - (N_T+O_T)は音声デ

ータNBにおける楽曲N全体の時間を、また $(N_E - N_T)$ は音声データNAにおける楽曲N全体の時間を示す。従って、第2項は音声データNBにおいて、楽曲Nの開始時点を基準とした演奏イベントの実行タイミングを示す。以上より、第1項と第2項の和、すなわちdは、音声データNBの最初のサンプル値の再生タイミングを基準とした演奏イベントの実行タイミングを示す。

【0153】

例えば、図16に例示されるSMFのデータが用いられる場合、最初の演奏イベントに関しては、 $D = 2. 11$ 、 $N_T = 1. 15$ 、 $N_E = 173. 11$ 、 $O_T = -0. 50$ 、 $O_E = 0. 40$ 、であるので、式(6)に従った計算により、この演奏イベントに対応する調整後のデルタタイムは約1.62秒となる。

【0154】

CPU62はステップS76における演奏イベントに関するデルタタイムの調整処理を終了すると、以下に説明するように、音声データNBとMIDIデータの再生を行う(ステップS77)。まず、CPU62はCDドライブ1に対し、音声データNBの再生命令を送信する。この再生命令に応じて、CDドライブ1は音声データNBのサンプル値を1/44100秒ごとに順次、コントローラ部6に送信する。CPU62は、CDドライブ1から音声データNBの最初のサンプル値、すなわちサンプル値(0)を受信すると、クロックから得られるクロック信号に基づき、その時点(以下、「基準時点R」と呼ぶ)を基準とする計時を開始する。

【0155】

その後、CPU62はサンプル値(0)に続き、サンプル値(1)、サンプル値(2)、・・・を順次受信する。CPU62は受信したサンプル値を順次、発音部4に送信する。発音部4はサンプル値を受信すると、これを音に変換し出力する。その結果、ユーザは楽曲Nを聞くことができる。

【0156】

CPU62は、サンプル値を発音部4に送信する処理を行うと同時に、RAM64に記録されている演奏イベントの調整後のデルタタイムと、基準時点Rを基準とする計時の結果とを順次比較し、計時の結果がデルタタイムと一致すると、

そのデルタタイムに対応する演奏イベントを自動演奏ピアノ部3に送信する。

【0157】

自動演奏ピアノ部3において、MIDIイベント制御回路34は演奏イベントをCPU62から受信すると、受信した演奏イベントを音源35もしくは駆動部36に送信する。音源35に演奏イベントが送信される場合、音源35は受信される演奏イベントに従い、楽器の音を示す音声データを順次、発音部4に送信する。発音部4は、既に再生が開始されている音声データNBの楽曲の音と共に、音源35から受信される楽器音による演奏をスピーカ44から出力する。一方、駆動部36に演奏イベントが送信される場合、駆動部36は受信される演奏イベントに従い、ピアノ31の鍵およびペダルを動かす。いずれの場合においても、ユーザは音声データNBに記録された楽曲Nと、SMFに記録された演奏情報による楽器音による演奏とを同時に聞くことができる。その後、CDドライブ1は、音声データNBの最後のサンプル値を送信し終えると音声データNBの再生を停止し、音声データNBとMIDIデータの再生が終了する。

【0158】

CPU62は、音声データNBの最後のサンプル値を受信し、ステップS77の処理を終了すると、操作表示部5に対し、トップオフセットおよびエンドオフセットの保存を行うか否かの指示をユーザに促すメッセージの表示命令を送信する。この表示命令に応じて、操作表示部5はメッセージを表示する（ステップS78）。ユーザがステップS78のメッセージに対し、操作表示部5の「保存する」に対応するキーパッドを押下すると、CPU62は操作表示部5から受信する信号に応じて、図16に示したSMFをRAM64から読み出し、読み出したSMFを保存命令と共にFDドライブ2に送信する。FDドライブ2は、SMFおよび保存命令を受信すると、FDに保存されているSMFを、新たに受信したSMFで上書き保存する（ステップS79）。CPU62は、FDドライブ2によるSMFの保存が終了すると、一連の演奏イベントのタイミング調整処理を終了する。一方、ユーザがステップS78のメッセージに対し、操作表示部5の「保存しない」に対応するキーパッドを押下すると、CPU62は操作表示部5から受信する信号に応じて、ステップS79の処理は行わず、一連の演奏イベント

のタイミング調整処理を終了する。

【0159】

上述した処理は、ステップS63およびステップS70における相関判定処理が共に成功した場合の処理であるが、音声データNAと音声データNBの内容が大きく異なる等の理由で、これらの相関判定処理が成功しない場合がある。以下、それらの場合における処理について説明する。

【0160】

まず、音声データNBの冒頭の20秒間分から得られる参照用加工済音声データ(i)の全てに関し、ステップS63における相関判定処理の結果が失敗であった場合、ステップS65の判定処理の結果がYesとなる。その場合、CPU62は、ユーザの手動による調整処理を実行する(ステップS80)。以下、図17を用いて、手動による調整処理を説明する。

【0161】

まずCPU62は、トップオフセットおよびエンドオフセットを記録する変数として、それぞれOTおよびOEを生成し、それらに0をセットする。続いて、CPU62はCDドライブ1に対し、音声データNBの再生の停止命令、および音声データNBの先頭からの再生命令を送信する。この再生命令に応じて、CDドライブ1は音声データNBのサンプル値を先頭から、1/44100秒ごとに順次、コントローラ部6に送信する。CPU62は、CDドライブ1から音声データNBの最初のサンプル値を受信すると、クロックから得られるクロック信号に基づき、その時点を基準とする計時を開始する。その後、CPU62は受信したサンプル値を順次、発音部4に送信する。その結果、ユーザは楽曲Nを聞くことができる。CPU62は、サンプル値を発音部4に送信すると同時に、RAM64に記録されているSMFの演奏イベントのデルタタイムと計時結果に従い、演奏イベントを自動演奏ピアノ部3に送信する。その結果、ユーザは楽曲Nと同時に、MIDIデータに従った楽器音による演奏を聞くことができる(ステップS91)。

【0162】

CPU62は、ステップS91における音声データNBおよびMIDIデータ

の再生を行う一方で、操作表示部5に対し、ユーザにトップオフセットの調整を促すメッセージの表示命令を送信する。ユーザは操作表示部5により表示されるメッセージに従い、楽曲Nに対しMIDIデータによる演奏が先行していると感じる場合には「-」に対応するキーパッドを、楽曲Nに対しMIDIデータによる演奏が遅れていると感じる場合には「+」に対応するキーパッドを押下する。CPU62は、「-」に対応するキーパッドが押下された場合には O_T を(1/75)だけ増加させ、「+」に対応するキーパッドが押下された場合には O_T を(1/75)だけ減少させる(ステップS92)。(1/75)は音楽CDにおける1フレームの時間(秒)を示す値である。

【0163】

CPU62は、ステップS92において O_T の値を更新すると、上述した式(6)に従い、RAM64のSMFに記録されている演奏イベントに対応するデルタタイムを用いて、調整後のデルタタイムを算出する。そして、調整後のデルタタイムをRAM64に記録する。以降、CPU62は、この調整後のデルタタイムに従い、演奏イベントの自動演奏ピアノ部3への送信を行う。その結果、MIDIデータによる演奏は、ユーザの操作に応じて全体として時間的に前後に調整される。従って、ユーザはトップオフセットの調整が反映されたMIDIデータによる演奏を楽曲Nと同時に聴くことにより、調整の是非を即座に確認できる。ユーザは「終了」に対応するキーパッドを押下するまで、繰り返しステップS92の調整操作を行うことができる(ステップS93:No)。

【0164】

ユーザは、トップオフセットの調整が反映されたMIDIデータによる演奏と楽曲Nとの同期が取れたと感じると、「終了」に対応するキーパッドを押下する。CPU62は、「終了」に対応するキーパッドが押下されると(ステップS93:Yes)、操作表示部5に対し、ユーザにエンドオフセットの調整を促すメッセージの表示命令を送信する。ユーザは操作表示部5により表示されるメッセージに従い、楽曲Nに対しMIDIデータによる演奏が徐々に早まっていると感じる場合には「-」に対応するキーパッドを、楽曲Nに対しMIDIデータによる演奏が徐々に遅れていると感じる場合には「+」に対応するキーパッドを押下

する。CPU62は、「-」に対応するキーパッドが押下された場合には O_E を(1/75)だけ増加させ、「+」に対応するキーパッドが押下された場合には O_E を(1/75)だけ減少させる(ステップS94)。

【0165】

CPU62は、ステップS94において O_E の値を更新すると、上述した式(6)に従い、RAM64のSMFに記録されている演奏イベントに対応するデルタタイムを用いて、調整後のデルタタイムを算出する。そして、調整後のデルタタイムをRAM64に記録する。CPU62はこの調整後のデルタタイムに従い、MIDIデータの再生を行う。その結果、MIDIデータによる演奏速度が、ユーザの操作に応じて調整され、ユーザはエンドオフセットの調整が反映されたMIDIデータによる演奏を楽曲Nと同時に聞くことにより、調整の是非を即座に確認できる。ユーザは「終了」に対応するキーパッドを押下するまで、繰り返しステップS94の調整操作を行うことができる(ステップS95:No)。

【0166】

ユーザは、エンドオフセットの調整が反映されたMIDIデータによる演奏と楽曲Nとの同期が取れたと感じると、「終了」に対応するキーパッドを押下する。CPU62は、「終了」に対応するキーパッドが押下されると(ステップS95:Yes)、手動による調整処理を終了する。

【0167】

手動による調整処理を終了した後、CPU62は図15のステップS76に処理を移し、手動により調整されたトップオフセットおよびエンドオフセットを用いて、ステップS76～ステップS79を実行する。その結果、ユーザは音声データNBによる楽曲Nと、楽曲Nに正しく同期の取れたMIDIデータの再生による演奏とを同時に聞くことができる。

【0168】

一方、音声データNBの楽曲Nの終了時点から約10秒手前から音声データNBの最後までの参照用加工済音声データ(i)の全てに関し、ステップS70における相関判定処理の結果が失敗であった場合、ステップS73の判定処理の結果がYesとなる。その場合、CPU62は、SMFに含まれる管理用イベント

による調整処理を実行する（ステップS81）。以下、管理用イベントによる調整処理を説明する。

【0169】

まずCPU62はCDドライブ1に対し、音声データNBの先頭からの再生命令を送信する。この再生命令に応じて、CDドライブ1は音声データNBのサンプル値を先頭から、1/44100秒ごとに順次、コントローラ部6に送信する。CPU62は、CDドライブ1から音声データNBの最初のサンプル値を受信すると、クロックから得られるクロック信号に基づき、その時点を基準とする計時を開始する。

【0170】

CPU62はCDドライブ1からサンプル値を受信すると、サンプル値を、サンプル値の受信時刻を示す時間情報とともにRAM64のキューに順次記録する。同時に、CPU62は受信したサンプル値の左右いずれかの絶対値が1000を超えるか否かを判定する。受信したサンプル値の左右いずれかの絶対値が1000を超えた場合、CPU62はDSP63に対し、管理用イベント生成処理の実行命令を送信し、DSP63は管理用イベント生成処理を実行する。管理用イベント生成処理は、図6を用いて既に説明した処理と同じであるので、説明を省略する。

【0171】

管理用イベント生成処理により、CPU62はDSP63から管理用イベントを受信する。CPU62は管理用イベントを受信すると、管理用イベントの受信時刻を示す時間情報と共に、RAM64に記録する。以下、RAM64に記録されているSMFに含まれる管理用イベントを管理用イベントAと呼び、音声データNBに対する管理用イベント生成処理によりCPU62が受信する管理用イベントを管理用イベントBと呼ぶ。

【0172】

CDドライブ1による音声データNBの再生が終了し、最後の管理用イベントBの時間情報をRAM64に記録すると、CPU62は管理用イベントAのデルタタイムと、管理用イベントBの時間情報を用いて、管理用イベントAと管理用

イベントBの対応付けの処理を行う。以下、データ例を用いて、管理用イベントAと管理用イベントBの対応付けの処理を説明する。

【0173】

図18は、管理用イベントAに対応するデルタタイムと、管理用イベントBに対応して記録される時間情報のデータ例を示す表である。ただし、音声データNAの列の先頭の値は先頭時間情報を、また音声データNBの列の先頭の値は先頭時間情報にステップS67においてRAM64に記録されたトップオフセットを加えた時間情報を示している。以下、先頭時間情報にトップオフセットを加えた時間情報を先頭時間情報Bと呼び、音声データNAに関する先頭時間情報を先頭時間情報Bと区別するために、先頭時間情報Aと呼ぶ。また、図18における第1の管理用イベントAを管理用イベントA₁、第2の管理用イベントAを管理用イベントA₂、…のように呼ぶ。管理用イベントBに関しても同様とする。

【0174】

図18において、先頭時間情報Aと先頭時間情報Bは対応している。まず、CPU62は(管理用イベントA₁−先頭時間情報A)／(管理用イベントB₁−先頭時間情報B)を求める。その値は、(1.51−1.15)／(1.01−0.65)=1.00である。続いて、CPU62は、求めた値が所定の値の範囲に入っているか否かを判定する。以下、例として、その範囲を0.97～1.03とする。この判定処理は、管理用イベントA₁と管理用イベントB₁が対応していると仮定した場合、先頭時間情報Aおよび先頭時間情報Bからの時間の誤差が3%以下であるか否かの判定処理である。この場合、誤差は0%であるので、CPU62は管理用イベントA₁と管理用イベントB₁を対応付ける。もし、上記の計算の結果が0.97より小さい場合、管理用イベントA₁のデルタタイムは管理用イベントB₁の時間情報に対して早すぎるので、CPU62は管理用イベントA₁に対応する管理用イベントBは無いものと見なし、管理用イベントA₂と管理用イベントB₁において、上記と同様の判定処理および対応付けの処理を行う。一方、もし上記の計算の結果が1.03より大きい場合、管理用イベントA₁のデルタタイムは管理用イベントB₁の時間情報に対して遅すぎるので、CPU62は管理用イベントB₁に対応する管理用イベントAは無いものと見なし、管

理用イベント A_1 と管理用イベント B_2 において、上記と同様の判定処理および対応付けの処理を行う。以下、最後に対応付けられた管理用イベント A と管理用イベント B を、それぞれ管理用イベント A_n 、管理用イベント B_n と呼ぶ。

【0175】

続いて、CPU62は、先頭時間情報 $B + (管理用イベント A_{n+1} - \text{先頭時間情報 } A) \times (管理用イベント B_n - \text{先頭時間情報 } B) / (管理用イベント A_n - \text{先頭時間情報 } A)$ を計算することにより、管理用イベント A_{n+1} のデルタタイムから、対応する管理用イベント B の時間情報を推定する。図19は、上記の計算式により算出される管理用イベント B の時間情報の推定値を示している。例えば、管理用イベント A_n が管理用イベント A_1 、管理用イベント B_n が管理用イベント B_1 である場合、管理用イベント B の推定値は、 $0.65 + (2.38 - 1.15) \times (1.01 - 0.65) / (1.51 - 1.15) = \text{約 } 1.88$ (秒) である。CPU62は、この推定値と、管理用イベント B_{n+1} との差が所定の範囲に入っているか否かを判定する。以下、例として、この範囲を $-0.20 \sim 0.20$ とする。この判定処理は、管理用イベント A_{n+1} のデルタタイムから推定された管理用イベント B の時間情報と、管理用イベント B_{n+1} の時間情報の示す時間の差が 0.20 秒以下であるか否かを判定する処理である。CPU62は、推定値と管理用イベント B_{n+1} との差が $-0.20 \sim 0.20$ の範囲に入っている場合、管理用イベント A_{n+1} と管理用イベント B_{n+1} とを対応付ける。

【0176】

一方、推定値から管理用イベント B_{n+1} の時間情報を減じた値が -0.20 より小さい場合、CPU62は管理用イベント A_{n+1} に対応する管理用イベント B は存在しないものと見なし、管理用イベント A_{n+1} の代わりに管理用イベント A_{n+2} を用いて、上記の判定処理および対応付けの処理を行う。また、推定値から管理用イベント B_{n+1} の時間情報を減じた値が 0.20 より大きい場合、CPU62は管理用イベント B_{n+1} に対応する管理用イベント A は存在しないものと見なし、管理用イベント B_{n+1} の代わりに管理用イベント B_{n+2} を用いて、上記の判定処理および対応付けの処理を行う。

【0177】

例えば、管理用イベントA_nが管理用イベントA₅であり、管理用イベントB_nが管理用イベントB₅である場合、管理用イベントA₆のデルタタイムから推定される管理用イベントBの時間情報は約8.25秒であるが、管理用イベントB₆の時間情報は9.76秒であり、その差は約-1.51秒となる。従って、管理用イベントA₆には対応する管理用イベントBが存在しないと見なされる。また、管理用イベントA_nが管理用イベントA₉であり、管理用イベントB_nが管理用イベントB₈である場合、管理用イベントA₁₀のデルタタイムから推定される管理用イベントBの時間情報は約17.79秒であるが、管理用イベントB₉の時間情報は15.57秒であり、その差は約2.22秒となる。従って、管理用イベントB₉には対応する管理用イベントAが存在しないと見なされる。

【0178】

CPU62は、全ての管理用イベントAおよび管理用イベントBに関し、上記の対応付けの処理を終えると、対応付けられた管理用イベントAのデルタタイムと管理用イベントBの時間情報との複数の対を用いて、管理用イベントAのデルタタイムと管理用イベントBの時間情報との間の関係式を推定する。図20は図19のデータ例を含むデータに関し、最小自乗法を用いて求めた回帰直線を示すグラフである。続いて、CPU62はRAM64に記録されたSMFから末尾時間情報を読み出し、読み出した末尾時間情報を推定した関係式に代入することにより、音声データNBにおける楽曲Nの最後に対応する時間情報を算出する。例えば、図20に示した例においては、(管理用イベントBの時間情報) = (管理用イベントAのデルタタイム) × 1.0053 - 0.5075の関係が推定されているので、末尾時間情報である173.11秒をこの式に代入することにより、約173.52秒が得られる。このようにして求められた音声データNBにおける楽曲Nの最後に対応する時間情報を、以下、末尾時間情報Bと呼び、音声データNAにおける末尾時間情報を末尾時間情報Bと区別するために、以下、末尾時間情報Aと呼ぶ。

【0179】

CPU62は、末尾時間情報Bを算出すると、(末尾情報時間B - 末尾情報時間A)を算出し、算出した値をエンドオフセットとしてRAM64に記録する。

続いて、CPU62はRAM64に記録されたSMFに、トップオフセットおよびエンドオフセットを含むシステムエクスクルーシブイベントを追加する。

その後、CPU62は図15のステップS76～ステップS79を実行する。その結果、ユーザは音声データNBによる楽曲Nと、楽曲Nに正しく同期の取れたMIDIデータの再生による演奏とを同時に聞くことができる。

【0180】

図21は、音声データNB、音声データNBから生成される参照用加工済音声データにおいて先頭参照用加工済音声データおよび末尾参照用加工済音声データに対応する部分、音声データNBから生成される管理用イベント、および演奏イベントの時間的関係を整理した図である。図21の下段に示されているように、演奏イベントはタイミング調整処理が行われない場合、音声データNBの開始時点から、2.11秒、2.62秒、3.60秒、・・・が経過した時点で、それぞれ自動演奏ピアノ部3に送信される。その場合、音声データNBの楽曲Nの開始前の沈黙は、音声データNAの楽曲Nの開始前の沈黙よりも短いため、MIDIデータによる演奏が遅れて始まる結果となる。また、音声データNBによる楽曲Nの全時間は、音声データNAによる楽曲Nの全時間よりも長いため、MIDIデータによる演奏速度は楽曲Nに対して速く、楽曲Nの後半においてはMIDIデータによる演奏が楽曲Nに先行する結果となる。これに対し、タイミング調整処理が行われた場合は、トップオフセットおよびエンドオフセットを用いた式(6)による演奏イベントのタイミング調整により、正しいタイミングで演奏イベントが自動演奏ピアノ部3に送信される。

【0181】

[1. 2. 2. 3: 同期再生および手動による再度の調整]

上述した演奏イベントのタイミング調整処理において、ステップS78でユーザがトップオフセットおよびエンドオフセットを含むSMFのFDへの保存を指示した場合、ユーザが再度、音声データNBとSMFとの同期再生を行う際には、図15を用いて説明した演奏イベントのタイミング調整処理の全てを再度実行する必要はない。以下、トップオフセットおよびエンドオフセットを含むSMFと音声データNBの同期再生における処理を説明する。

【0182】

ユーザが操作表示部5の所定のキーパッドを押下することにより、FDに保存されているSMFと音楽CD-Bに含まれる音声データNBとの同期再生を指示すると、まず、CPU62はSMFを、FDドライブ2を介してFDから読み出し、読み出したSMFをRAM64に記録する。続いて、CPU62はSMFのトップオフセットおよびエンドオフセットを読み出し、SMFに含まれる演奏イベントのデルタタイムに対し、式(6)を用いた調整処理を行った後、RAM64に調整処理後のデルタタイムを記録する。

【0183】

続いて、CPU62はCDドライブ1に音声データNBの再生命令を送信し、CDドライブ1は音声データNBのサンプル値を順次、コントローラ部6に送信する。CPU62は、最初のサンプル値を受信すると、そのサンプル値の受信時刻を基準として、計時を開始する。CPU62は、受信するサンプル値を順次、発音部4に送信し、ユーザは楽曲Nを聞くことができる。同時に、CPU62は調整処理後のデルタタイムと計時される時間とを順次比較し、調整処理後のデルタタイムと計時される時間が一致した場合、そのデルタタイムに対応する演奏イベントを自動演奏ピアノ部3に送信する。その結果、ユーザはSMFのMIDIデータによる演奏を、楽曲Nに対する正しいタイミングで聞くことができる。

【0184】

ユーザが楽曲NおよびMIDIデータによる演奏を同時に聴いている際、それらが正しく合っていないと感じた場合、ユーザは操作表示部5の所定のキーパッドを押下することにより、手動によるトップオフセットおよびエンドオフセットの調整処理を行うことができる。手動によるトップオフセットおよびエンドオフセットの調整処理は、図17のステップS91～ステップS95までの処理と同様であるので、説明を省略する。ユーザは手動によるトップオフセットおよびエンドオフセットの調整処理を終えた後、操作表示部5の所定のキーパッドを押下することにより、調整後のトップオフセットおよびエンドオフセットを含むSMFをFDに保存することができる。CPU62は、操作表示部5を介して、ユーザよりSMFのFDへの保存の指示を受信すると、調整後のトップオフセットお

およびエンドオフセットを含むSMFをRAM64から読み出し、読み出したSMFを保存命令と共にFDドライブ2に送信する。FDドライブ2はSMFを受信し、FDに保存されているSMFに対し、受信したSMFを上書き保存する。その結果、再度ユーザが音声データNBとSMFのMIDIデータとの同期再生を行う際、最終的に調整の行われたトップオフセットおよびエンドオフセットを用いて演奏イベントの送信タイミングが調整され、再度の手動による調整処理を要しない。

【0185】

[2:変形例]

上述した実施形態は本発明の実施形態の例示であり、上記実施形態に対しては、本発明の趣旨から逸脱しない範囲で様々な変形を加えることができる。以下、変形例を示す。

【0186】

[2. 1: 第1変形例]

上述した実施形態においては、SMFに記録される音声データNAに対する管理用イベントおよび演奏イベントのデルタタイム、先頭時間情報、および末尾時間情報は、CPU62が音声データNAの最初のサンプル値を受信した時点を基準とする、CPU62の計時により得られる時間情報であった。同様に、上述した実施形態においては、RAM64に記録される音声データNBに対する管理用イベントのデルタタイム、音声データNBにおいて楽曲Nの開始に対応する時間情報、および音声データNBにおいて楽曲Nの最後に対応する時間情報は、CPU62が音声データNBの最初のサンプル値を受信した時点を基準とする、CPU62の計時により得られる時間情報であった。

【0187】

これに対し、第1変形例においては、CDドライブ1からコントローラ部6に送信される、音声データNAおよび音声データNBのタイムコードが、CPU62の計時による時間情報の代わりに用いられる。タイムコードとは、音楽CDにおいて音声データのまとまりである1フレーム、すなわち588のサンプル値に対応して音楽CDに記録されているデータであり、各タイムコードは、音声デー

タの開始時点から当該タイムコードに対応した音声データの再生タイミングまでの経過時間を表している。

【0188】

第1変形例において、まずコントローラ部6はクロック信号を常時、CDドライブ1に送信している。CDドライブ1は、コントローラ部6から受信するクロック信号に従い、音声データのコントローラ部6への送信を行う。さらに、CDドライブ1は、音声データの送信に際し、音声データに対応するタイムコードも送信する。CPU62はRAM64のキューにサンプル値を記録する際、計時による時間情報ではなく、CDドライブ1から送信されるタイムコードの示す時間情報をサンプル値と共に記録する。なお、より細かい精度で時間情報を扱う必要がある場合には、CPU62はタイムコードと、そのタイムコードに対応するフレームの先頭から数えた各サンプル値の番号から、各サンプル値に対応するより正確な時間情報を算出し、算出した時間情報をキューに記録する。

【0189】

CPU62は、音声データNAに対する管理用イベントや演奏イベント、および音声データNBに対する管理用イベントを受信すると、その時点で受信されたサンプル値に対応するタイムコードが示す時間情報を、それらのイベントデータに対応する時間情報をとする。

第1変形例によれば、音楽CDに記録されているタイムコードが利用され、CPU62が計時を行う必要がないため、CPU62の処理が簡便化される。

【0190】

[2. 2 : 第2変形例]

第2変形例においては、同期記録再生装置SSの各構成部は同じ装置の中に配置されておらず、グループごとに分離して配置されている。例えば、以下のそれぞれのグループに分離配置することが可能である。

- (1) CDドライブ1
- (2) FDドライブ2
- (3) 自動演奏ピアノ部3
- (4) ミキサ41およびD/Aコンバータ42

- (5) アンプ4 3
- (6) スピーカ4 4
- (7) 操作表示部5 およびコントローラ部6

【0191】

さらに、コントローラ部6は、記録動作のみを行う装置と再生動作のみを行う装置とに別々に構成されていてもよい。これらの構成部のグループ間は、オーディオケーブル、MIDIケーブル、オーディオ用光ケーブル、USB (Universal Serial Bus) ケーブル、および専用の制御用ケーブル等で接続される。また、FDドライブ2、アンプ4 3、スピーカ4 4等は市販のものを利用してもよい。

【0192】

第2変形例によれば、同期記録再生装置SSの配置の柔軟性が高まると同時に、ユーザは同期記録再生装置SSの全てを新たに準備することなく、必要なコストを低減できる。

【0193】

[2. 3 : 第3変形例]

第3変形例においては、同期記録再生装置SSにおいてCDドライブ1およびFDドライブ2はない。その一方で、通信インターフェースはLAN (Local Area Network) に接続可能な機能を有し、外部の通信機器とLANおよびWANを介して接続されている。さらに、コントローラ部6はHD (Hard Disk) を有している。

【0194】

コントローラ部6は、LANを介して他の通信機器から、音声データを受信し、受信した音声データをHDに記録する。同様に、コントローラ部6は、LANを介して他の通信機器から、音声データに対応して作成されたSMFを受信し、受信したSMFをHDに記録する。

【0195】

コントローラ部6はCDドライブ1から音声データを受信する代わりに、HDから音声データを読み出す。また、コントローラ部6はFDドライブ2に対しS

MFの書込および読出を行う代わりに、HDに対し同様の動作を行う。

【0196】

第3変形例によれば、ユーザは音声データおよびSMFを、LANを介して地理的に離れた通信機器との間で送受信することができる。なお、LANはインターネット等の広域通信網と接続されていてもよい。

【0197】

[2.4: 第4変形例]

上述した実施形態においては、相関判定処理のステップS51およびステップS52において、絶対相関指標による判定、相対相関指標による判定、および相関値による判定の全てが用いられているが、第4変形例においては、これらの判定の1つまたは複数の組み合わせにより相関判定処理が行われる。なお、これらの判定の1つまたは複数の組み合わせを、自由に選択可能としてもよい。

第4変形例によれば、より柔軟に必要とされる精度の判定結果を得ることができる。

【0198】

[2.5: 第5変形例]

上述した実施形態においては、相関判定処理のステップS52において、式(4)および式(5)で示される判定により、相関値の極大値が検出されているが、第5変形例においては、式(4)で示される判定のみが行われ、相関値の極値が検出される。

【0199】

より具体的には、ステップS52において、DSP63は D_{m-1} と D_m の積を求め、その積が0以下であるか否かを判定する。ここでその積が0以下である場合、相関値の変化率は0であるか、もしくは0をまたがって変化しているので、その時点における相関値は極値もしくは極値の近似値である。従って、 D_{m-1} と D_m の積が0以下である場合、ステップS52の判定処理の結果をYesとする。

【0200】

第5変形例によれば、極大値の近辺に極小値が現れる可能性が低い場合、上述した実施形態におけるステップS52と同様の判定結果を、より簡易な判定処理

により得ることができる。

【0201】

[2. 6: 第6変形例]

上述した実施形態においては、SMFに記録される先頭時間情報および末尾時間情報は、音声データNAのサンプル値の値が一定値を超えるか否かの判定に基づき決定されていた。従って、SMFに記録される先頭参照用加工済音声データおよび末尾参照用加工済音声データは、それぞれ、音声データNAにおいて楽曲Nの先頭付近および末尾付近に位置する音声データを用いて生成されるデータであった。

【0202】

しかしながら、第6変形例においては、先頭時間情報および末尾時間情報はそれぞれ、楽曲Nの任意の時点が選択される。例えば、ユーザが楽曲Nの内容に応じて適する時点をそれらの時間情報として指定することができる。また、音声データNAの先頭から一定時間経過後の時点を先頭時間情報とし、また、音声データNAの末尾から一定時間手前の時点を末尾時間情報とすることもできる。

【0203】

第6変形例によれば、例えばライブ演奏を収録した音楽CDにおいて、楽曲の前後の拍手や歓声の部分を避けて、先頭参照用加工済音声データあるいは末尾参照用加工済音声データの生成用データを選択することが可能である。また、楽曲Nの先頭付近あるいは末尾付近に特定のパターンの繰り返しがある場合であっても、そのパターン部分を避けて、相関判定処理に用いるのに適した特徴的な部分の音声データを先頭参照用加工済音声データあるいは末尾参照用加工済音声データの生成用データとして選択することが可能である。

【0204】

[2. 7: 第7変形例]

第7変形例においては、トップオフセットおよびエンドオフセットを含むSMFがFDに記録される際に、音声データNBを特定可能な識別情報がSMFに追加される。音声データNBを特定可能な識別情報としては、例えば音楽CD-Bの識別データおよび音楽CD-Bにおいて音声データNBが記録されているトラ

ック番号の組み合わせが考えられる。なお、音楽CD-Bの識別データは音楽CD-Bの目次情報として記録されており、CPU62はCDドライブ1を介して識別データを音楽CD-Bから読み出す。CPU62は、読み出した音楽CD-Bの識別データを、ユーザの指示により与えられている音声データNBのトラック番号と共に、RAM64に記録されているSMFにシステムエクスクルーシブイベントとして追加する。その後、CPU62はRAM64に記録されているSMFをFDドライブ2に送信し、FDドライブはそのSMFをFDに保存する。

【0205】

ユーザが音楽CDをCDドライブ1にセットし、再生を行う音声データが含まれるトラック番号を操作表示部5を用いて指定すると、CPU62はまず、CDドライブ1を介して音楽CDの識別データを読み出し、続いて読み出した音楽CDの識別データとユーザにより指定されたトラック番号の組み合わせが記録されているSMFを、FDドライブ2を介して、FDに記録されている複数のSMFの中から検索する。CPU62は検索されたSMFを読み出し、読み出したSMFを用いて音声データとMIDIデータとの同期再生を開始する。SMFの検索において、CDドライブ1にセットされた音楽CDの識別データとユーザにより指定されたトラック番号の組み合わせを記録したSMFがFDに記録されていない場合、CPU62は音声データとMIDIデータとの同期再生を行わず、操作表示部5を介してエラーメッセージを表示し、ユーザに正しいFDもしくは音楽CDのセットを促す。

【0206】

第7変形例によれば、ユーザが正しくない組み合わせの音楽CDとSMFを用いて同期再生を試みた場合、組み合わせが正しくないことが即座にユーザに対して示されるので、音声データとSMFの管理を容易に行うことができる。さらに、FDに複数のSMFが記録されている場合であっても、セットされた音楽CDと指定されたトラック番号に基づき、自動的に適するSMFが読み出されるため、ユーザはSMFの指定を別途行う必要がない。

【0207】

【発明の効果】

本発明によれば、同じ楽曲であっても、楽曲の開始時点および終了時点が異なる音声データを記録した、異なる版の音声データのいずれに対しても、正しいタイミングでMIDIデータ等の演奏データの同期再生を行うことができる。従って、同じ楽曲の異なる版に対し異なる演奏データを準備する必要がなく、演奏データの作成および管理が簡便化される。

【0208】

なお、同じ楽曲の異なる版においては、楽曲の録音レベルが異なる場合があるが、本発明によれば、楽曲の開始時点および終了時点を決定する際に用いられる指標の一つとして、参照用の音声データの示す音声波形の形状と、実際の音声データの示す音声波形の形状との相似の度合いを示す指標を用いることが可能なため、録音レベルが異なる版のオーディオデータに対しても、正しく再生開始のタイミングを決定することができる。

【図面の簡単な説明】

【図1】 本発明の実施形態に係る同期記録再生装置SSの構成を示す図である。

【図2】 SMFの構成を示す図である。

【図3】 MIDIイベントの構成を示す図である。

【図4】 本発明の実施形態に係る相関判定用データ生成処理のフロー図である。

【図5】 本発明の実施形態に係る相関判定用データ生成処理に用いられる樹形フィルタの構成を示す図である。

【図6】 本発明の実施形態に係る管理用イベント生成処理のフロー図である。

【図7】 本発明の実施形態に係る管理用イベント生成処理に用いられる樹形フィルタの構成を示す図である。

【図8】 本発明の実施形態に係る中期指標、長期指標および管理用イベントの時間的関係を示す図である。

【図9】 本発明の実施形態に係る管理用イベントと演奏イベントの時間的関係を示す図である。

【図10】 本発明の実施形態に係る末尾参照用加工済音声データ生成処理のフロー図である。

【図11】 本発明の実施形態に係る相関判定処理のフロー図である。

【図12】 本発明の実施形態に係る相関判定処理における、計算式の値の変化と判定結果との関係を示す図である。

【図13】 本発明の実施形態に係るSMFを示す図である。

【図14】 本発明の実施形態に係る音声データ、先頭参照用加工済音声データ、末尾参照用加工済音声データ、管理用イベントおよび演奏イベントの時間情報およびデルタタイムの時間的関係を示す図である。

【図15】 本発明の実施形態に係る演奏イベントのタイミング調整処理のフロー図である。

【図16】 本発明の実施形態に係るSMFを示す図である。

【図17】 本発明の実施形態に係る手動による調整処理のフロー図である。

【図18】 本発明の実施形態に係る、異なる音声データに対する管理用イベントのデルタタイムおよび時間情報のデータ例である。

【図19】 本発明の実施形態に係る管理用イベントの対応付け処理の様子を示すデータ例である。

【図20】 本発明の実施形態に係る、異なる音声データに対する管理用イベントのデルタタイムおよび時間情報の関係を示す図である。

【図21】 本発明の実施形態に係る音声データ、参照用加工済音声データ、管理用イベントおよび演奏イベントの時間情報およびデルタタイムの時間的関係を示す図である。

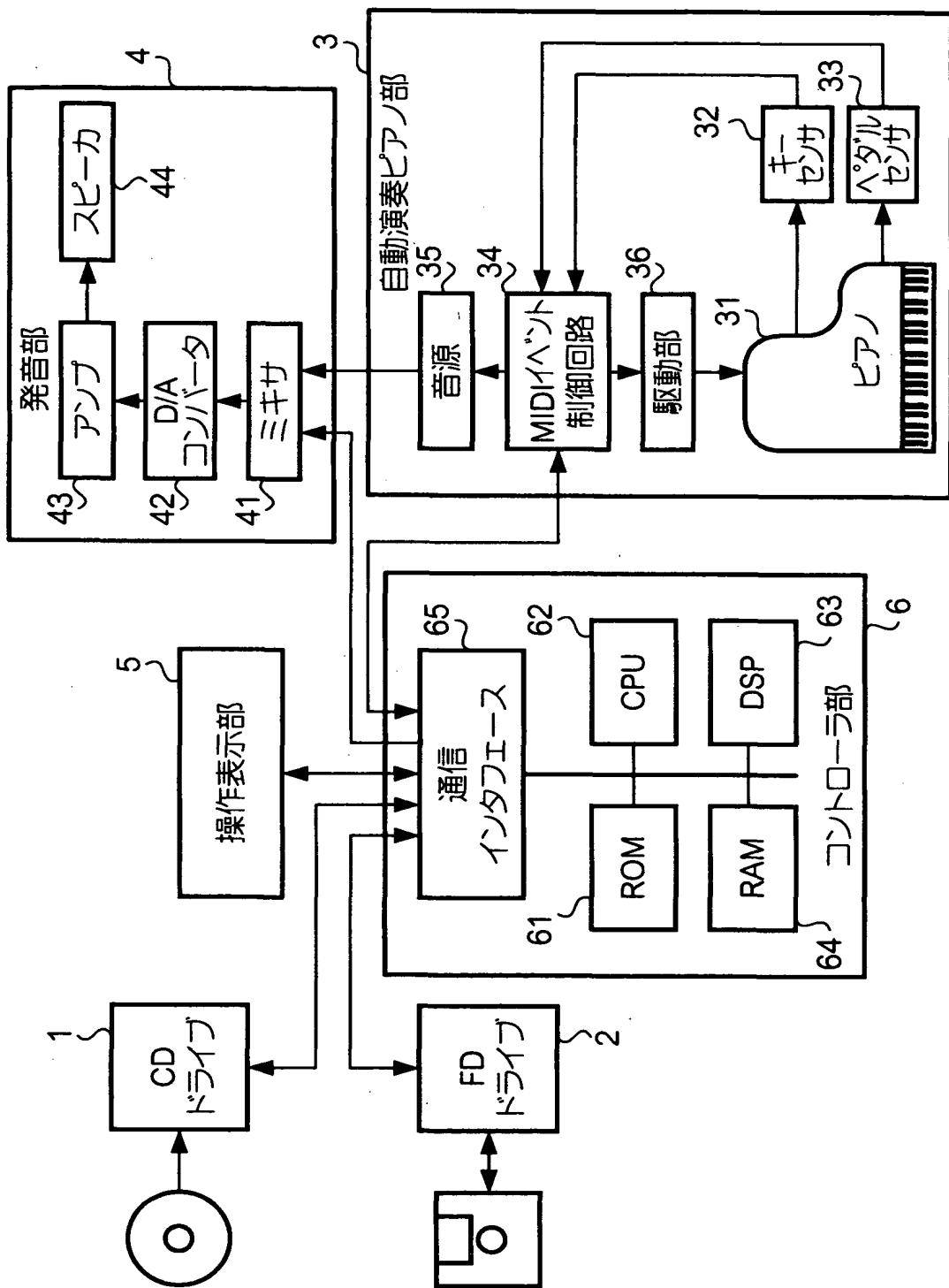
【符号の説明】

1 . . . CDドライブ、2 . . . FDドライブ、3 . . . 自動演奏ピアノ部、4 . . . 発音部、5 . . . 操作表示部、6 . . . コントローラ部、31 . . . ピアノ、32 . . . キーセンサ、33 . . . ペダルセンサ、34 . . . M I D I イベント制御回路、35 . . . 音源、36 . . . 駆動部、41 . . . ミキサ、42 . . . D/Aコンバータ、43 . . . アンプ、44 . . . スピーカ、61 . . . R

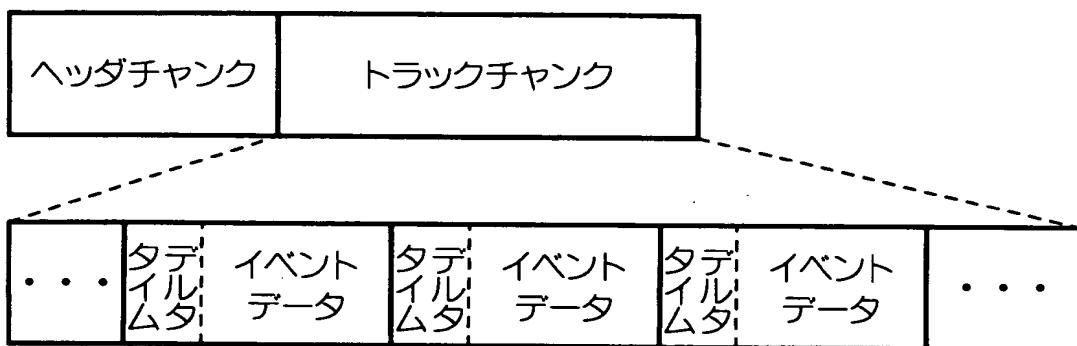
OM、62・・・CPU、63・・・DSP、64・・・RAM、65・・・通信インターフェース。

【書類名】図面

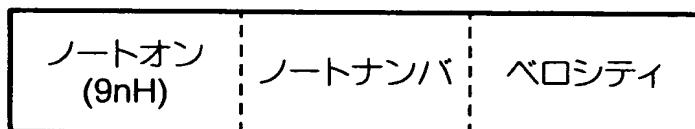
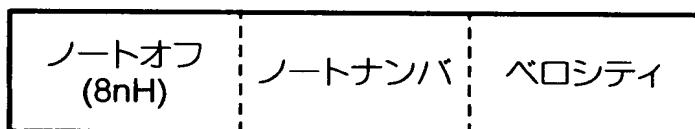
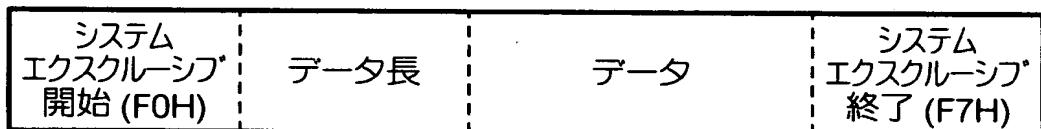
【図1】



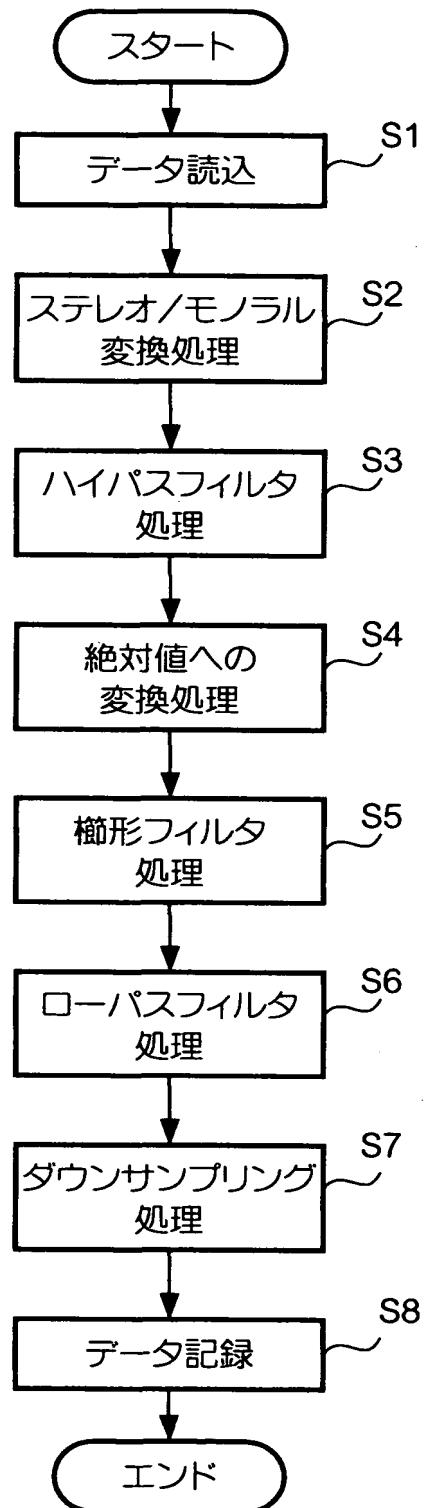
【図2】



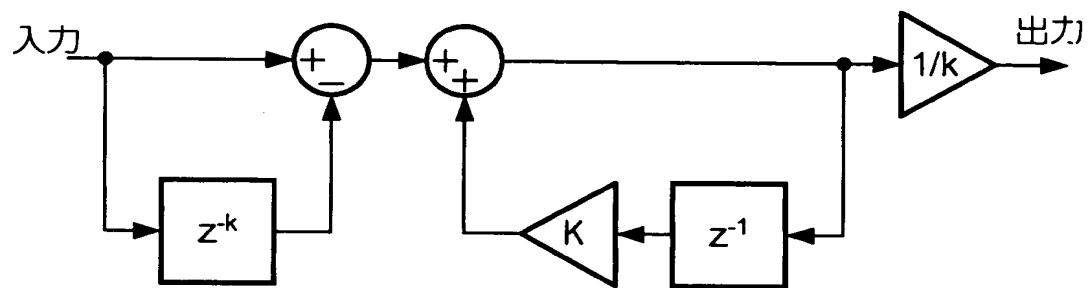
【図3】

ノートオンイベントノートオフイベントシステムエクスクルーシブイベント

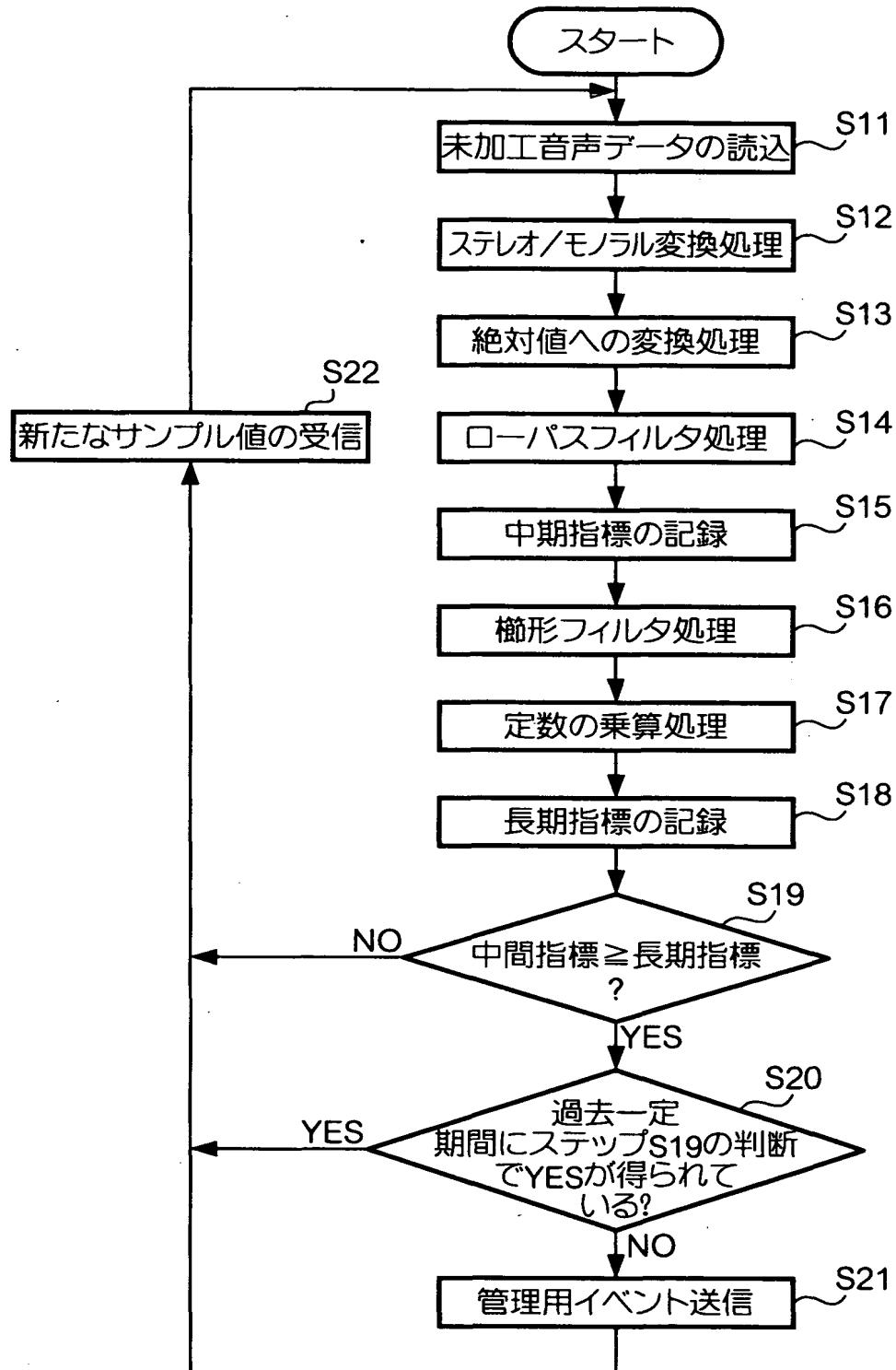
【図4】



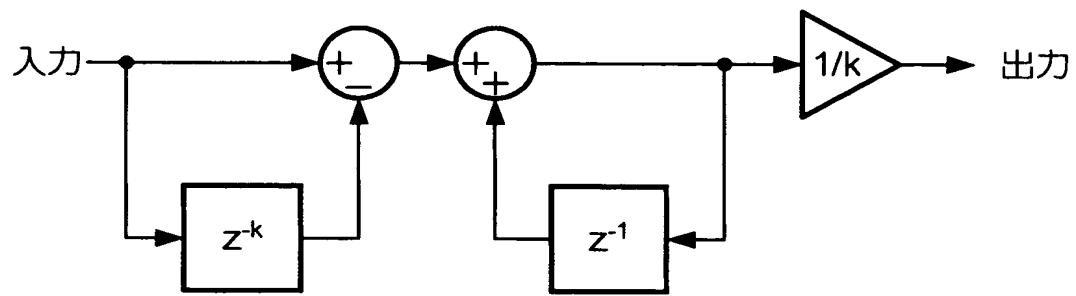
【図5】



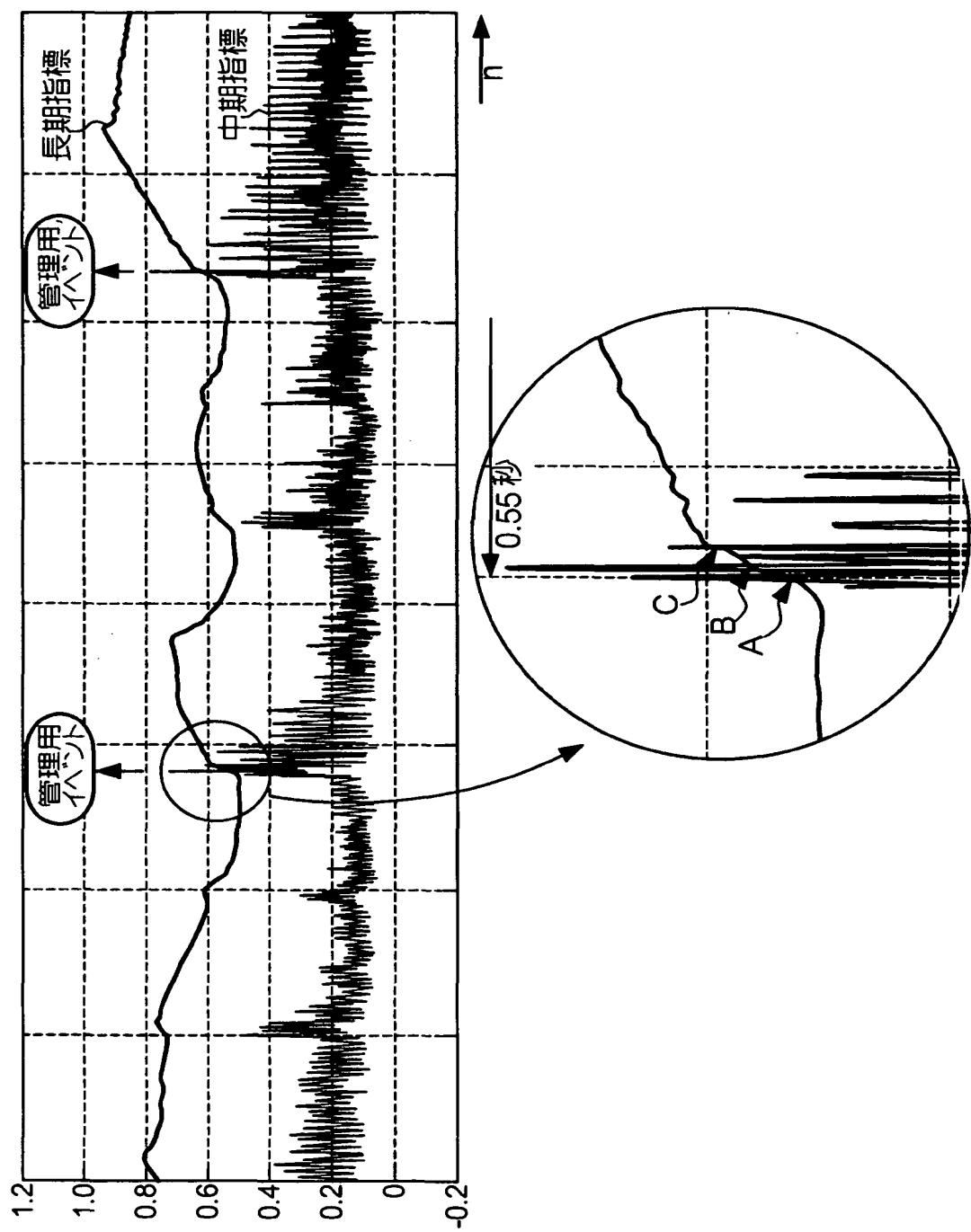
【図6】



【図7】

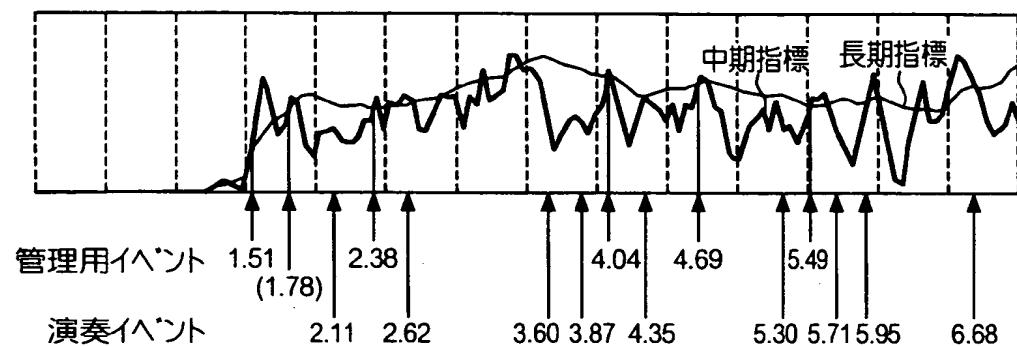


【図8】

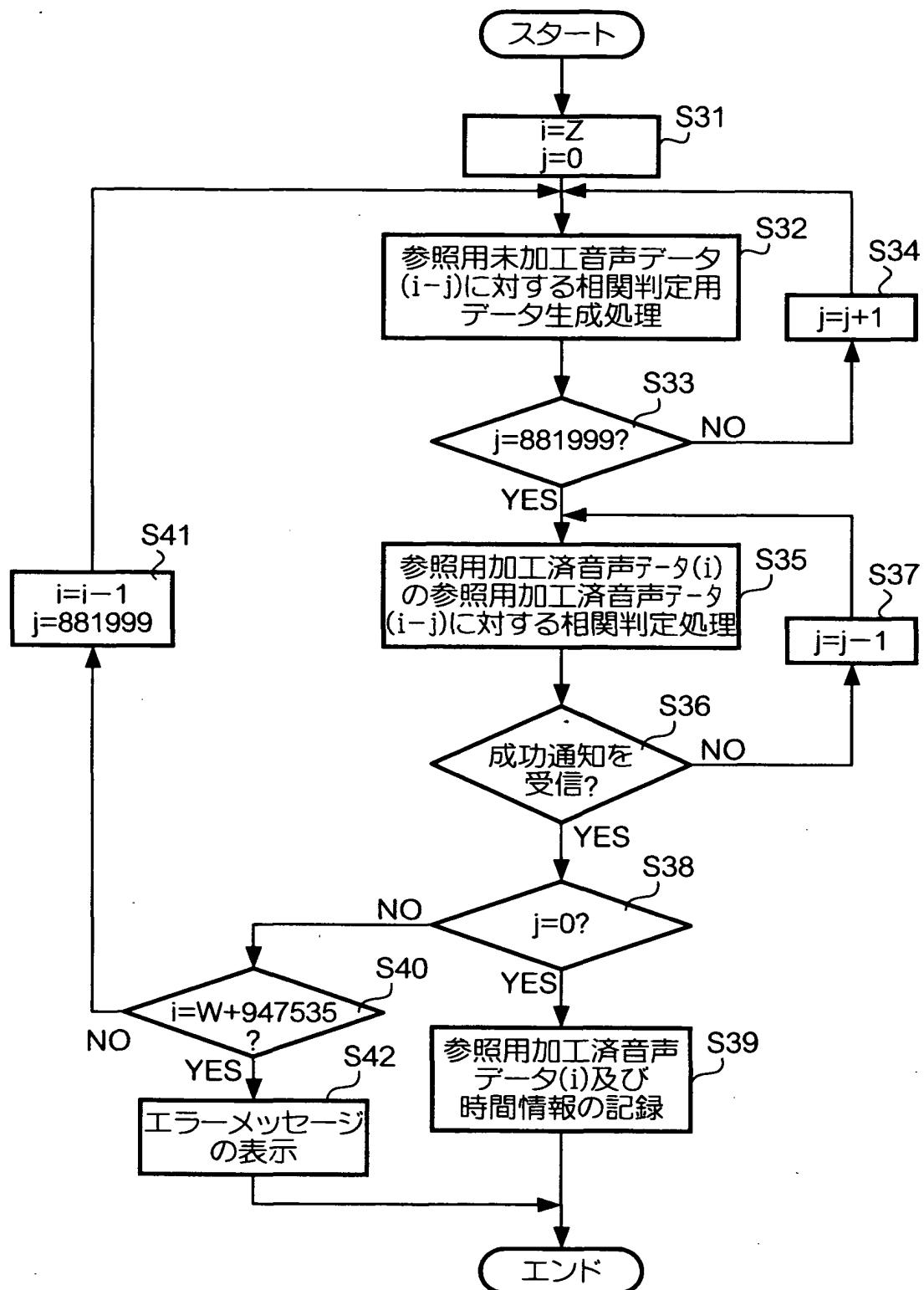


【図9】

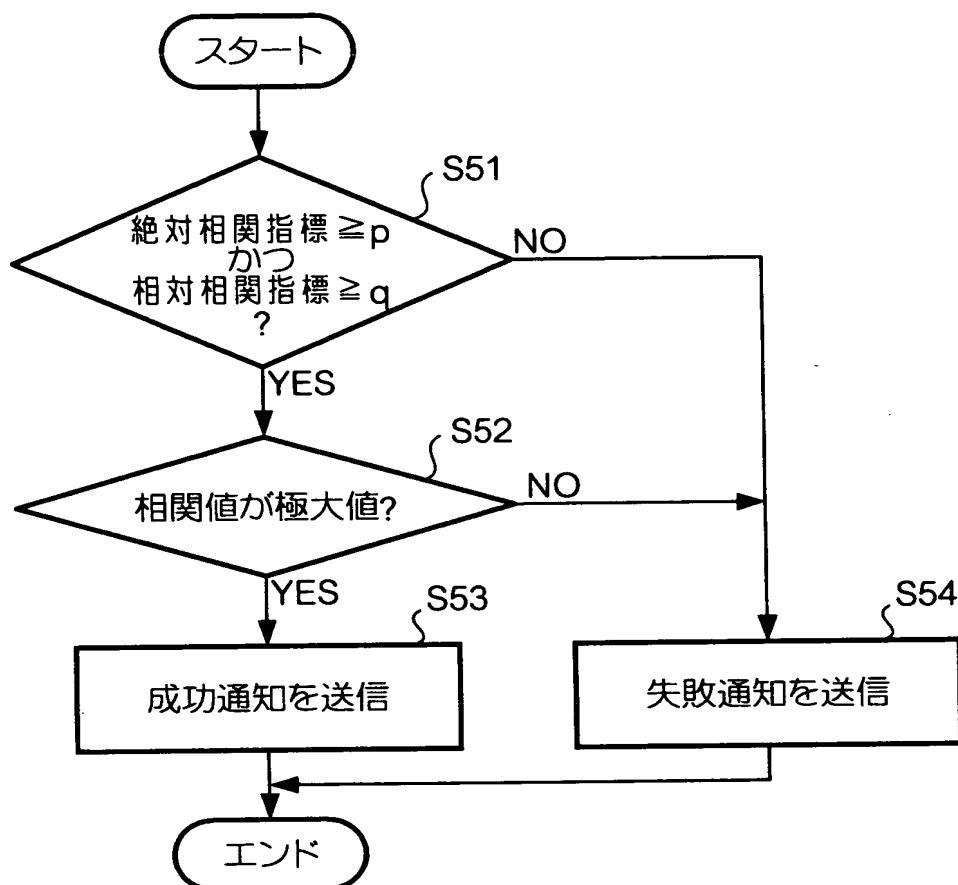
音声データNA



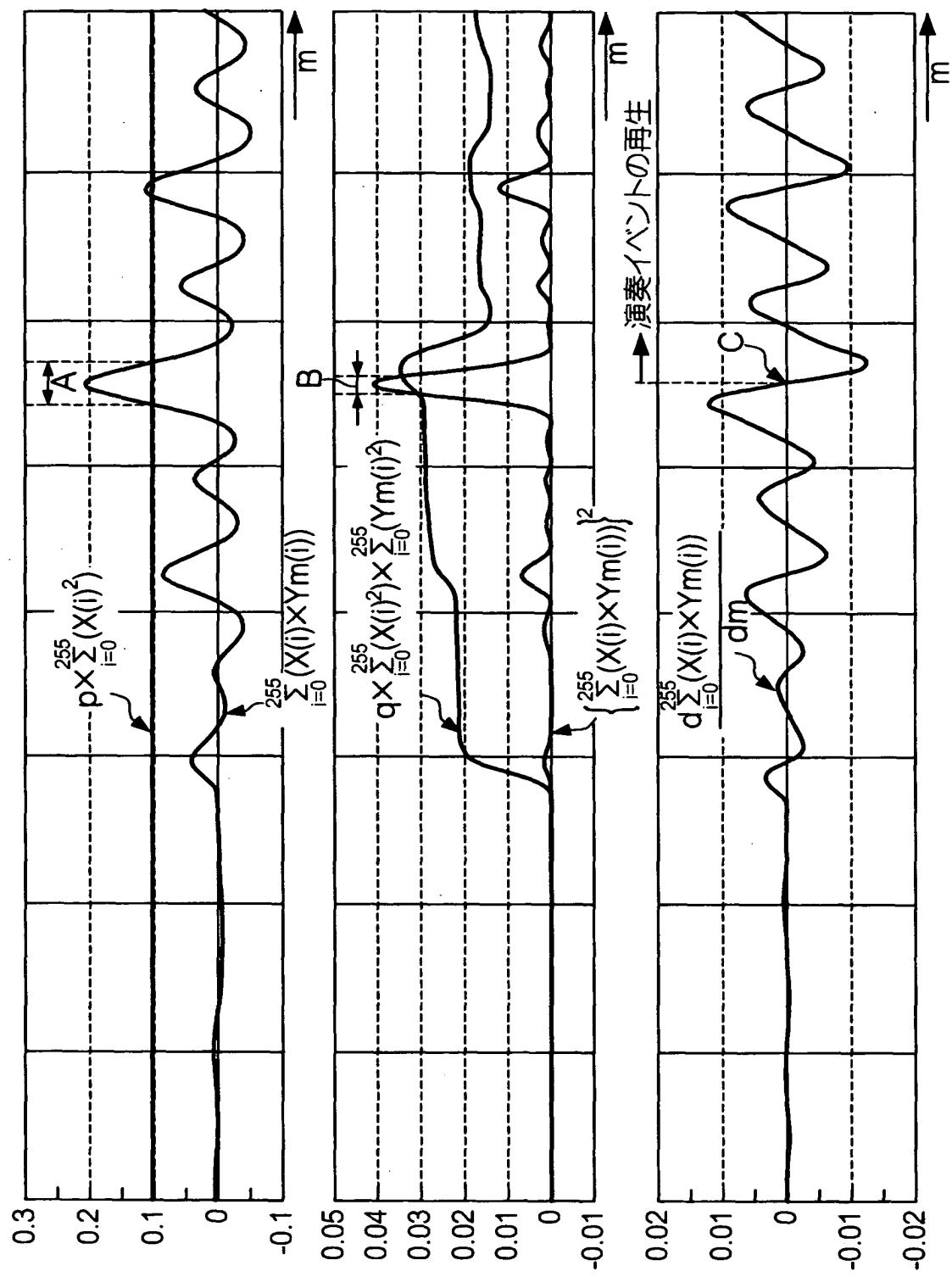
【図10】



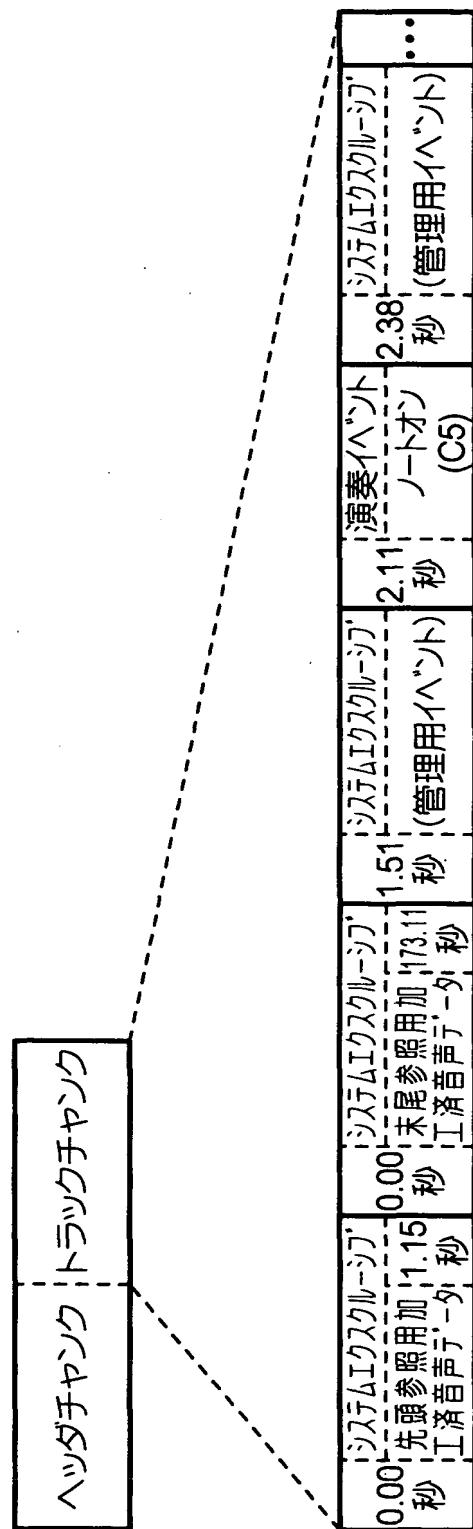
【図11】



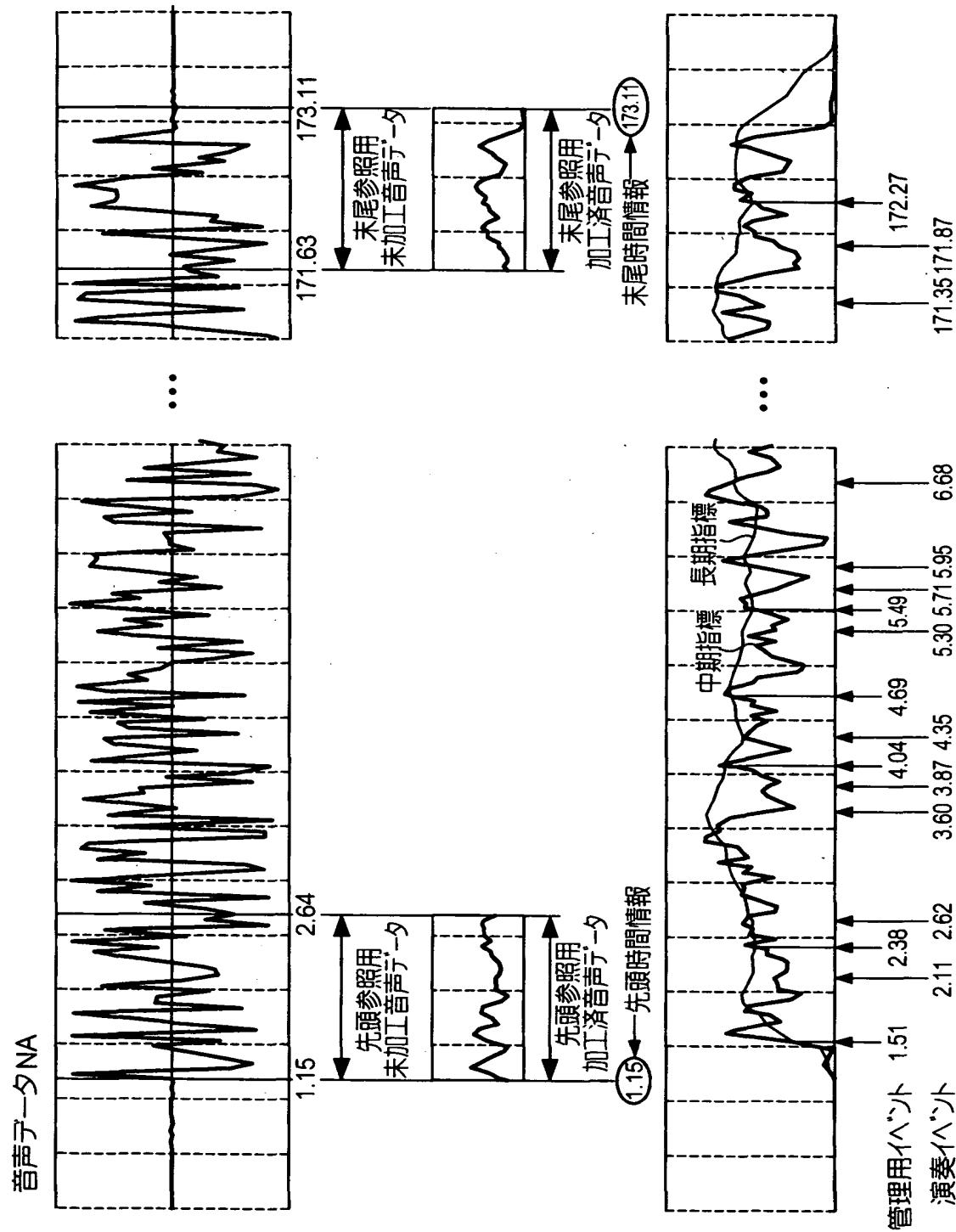
【図12】



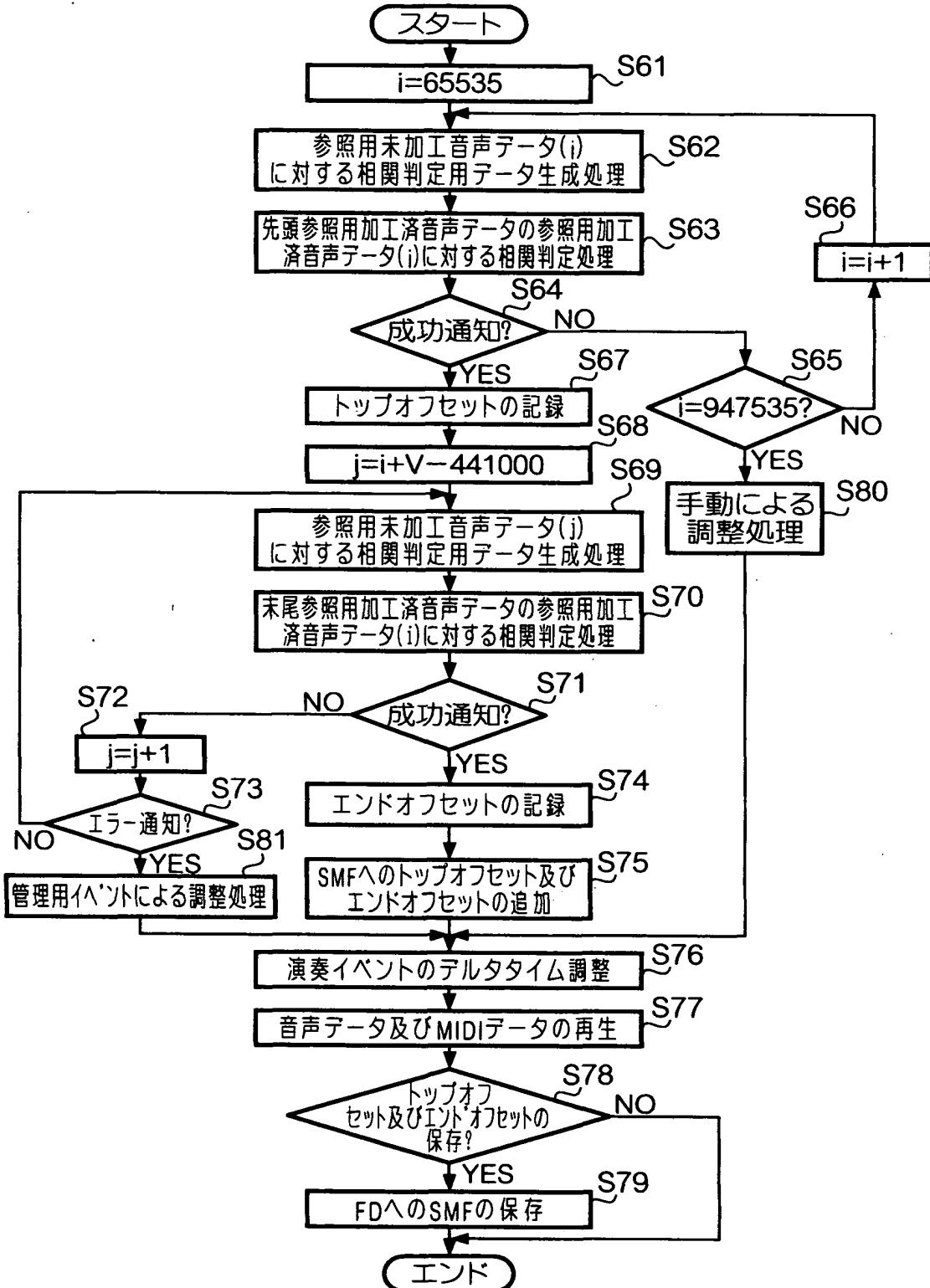
【図13】



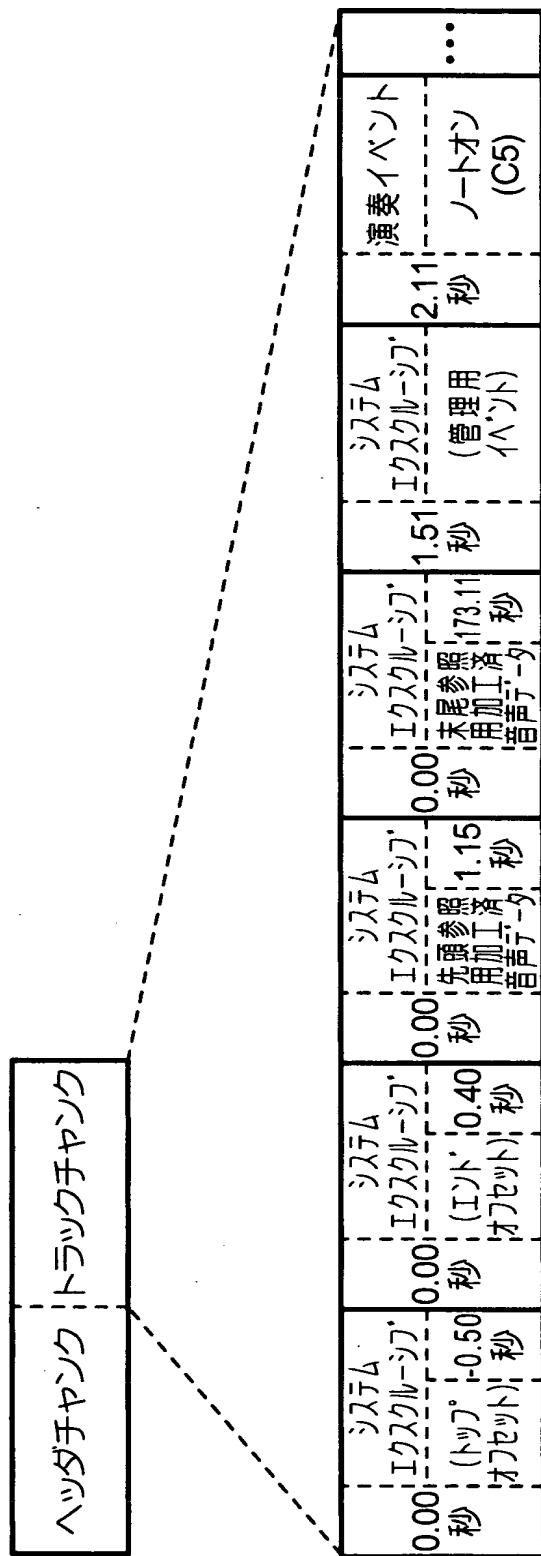
【図14】



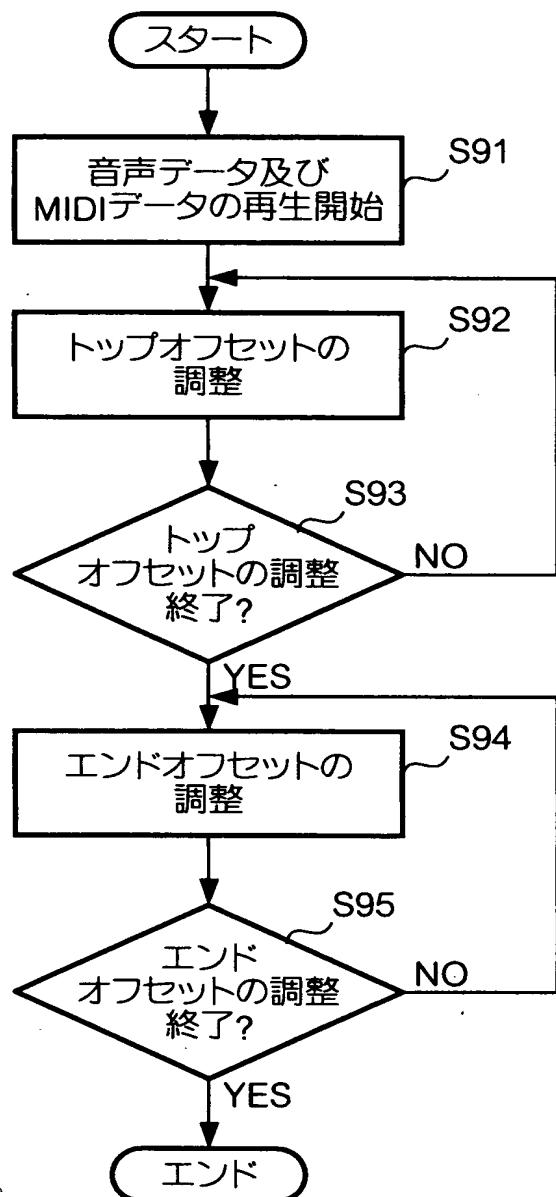
【図15】



【図16】



【図17】



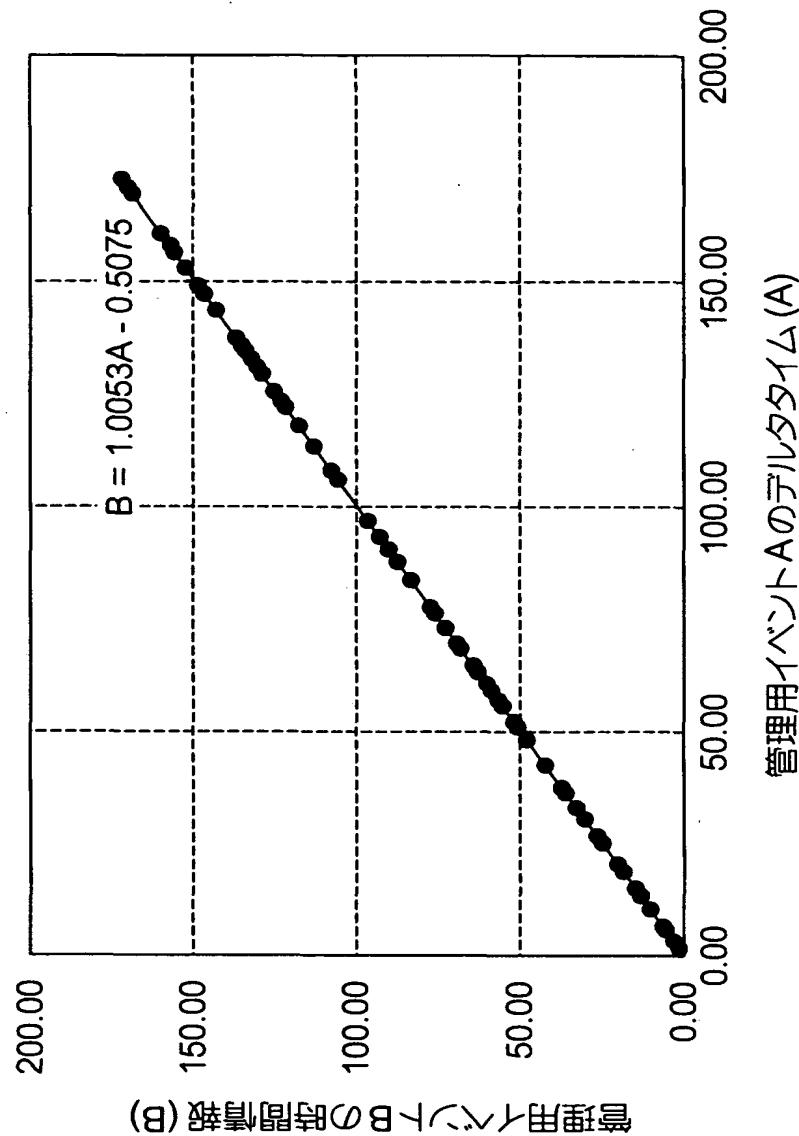
【図18】

音声データNA (SMF)			音声データNB		
		デルタ タイム(秒)			時間 情報(秒)
0	先頭時間情報	1.15	0	先頭時間情報 +トップオフセット	0.65
1	管理用イベントA	1.51	1	管理用イベントB	1.01
2	管理用イベントA	2.38	2	管理用イベントB	1.88
3	管理用イベントA	4.04	3	管理用イベントB	3.54
4	管理用イベントA	4.69	4	管理用イベントB	4.19
5	管理用イベントA	5.49	5	管理用イベントB	4.98
6	管理用イベントA	8.77	6	管理用イベントB	9.76
7	管理用イベントA	10.19	7	管理用イベントB	12.34
8	管理用イベントA	12.77	8	管理用イベントB	13.91
9	管理用イベントA	14.35	9	管理用イベントB	15.57
10	管理用イベントA	18.22	10	管理用イベントB	17.78
11	管理用イベントA	19.84	11	管理用イベントB	19.43
12	管理用イベントA	22.90	12	管理用イベントB	24.45
⋮	⋮	⋮	⋮	⋮	⋮
69	管理用イベントA	165.47	69	管理用イベントB	160.67
70	管理用イベントA	169.94	70	管理用イベントB	168.49
71	管理用イベントA	170.77	71	管理用イベントB	170.32
72	管理用イベントA	172.27	72	管理用イベントB	171.13
			73	管理用イベントB	171.85
			74	管理用イベントB	172.67

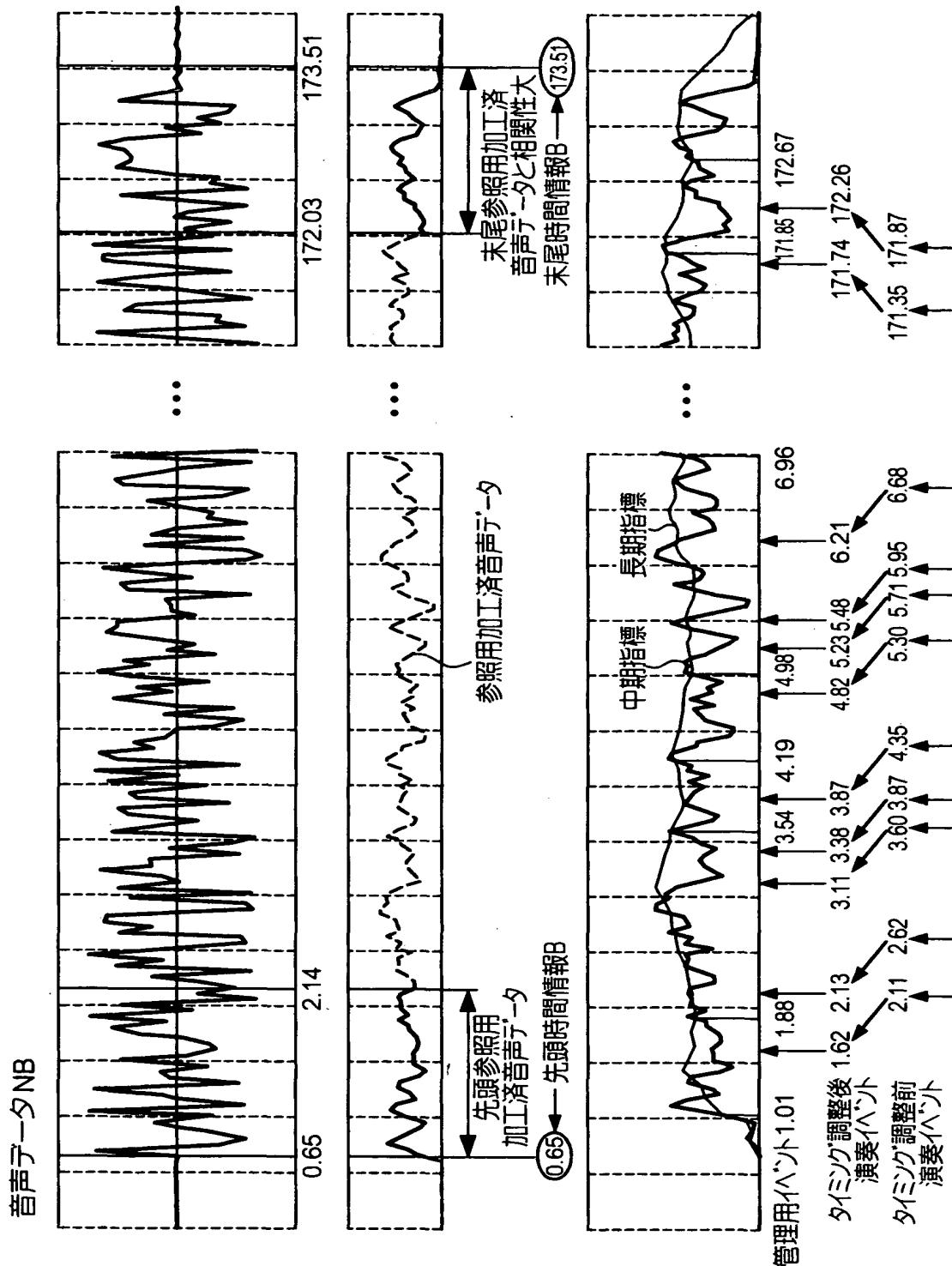
【図19】

音声データNA (SMF)			管理イベントBの時間情報の推定値	音声データNB		
		テルタ タイム (秒)				時間 情報(秒)
0	先頭時間情報	1.15	—	0	先頭時間情報 +トップオフセット	0.65
1	管理用イベントA	1.51	—	1	管理用イベントB	1.01
2	管理用イベントA	2.38	1.88	2	管理用イベントB	1.88
3	管理用イベントA	4.04	3.54	3	管理用イベントB	3.54
4	管理用イベントA	4.69	4.19	4	管理用イベントB	4.19
5	管理用イベントA	5.49	4.99	5	管理用イベントB	4.98
6	管理用イベントA	8.77	8.25	—	—	—
7	管理用イベントA	10.19	9.67	6	管理用イベントB	9.76
8	管理用イベントA	12.77	12.35	7	管理用イベントB	12.34
9	管理用イベントA	14.35	13.93	8	管理用イベントB	13.91
—	—	—	—	9	管理用イベントB	15.57
10	管理用イベントA	18.22	17.79	10	管理用イベントB	17.78
11	管理用イベントA	19.84	19.41	11	管理用イベントB	19.43
12	管理用イベントA	22.90	22.51	—	—	—
13	管理用イベントA	24.84	24.46	12	管理用イベントB	24.45
⋮	⋮	⋮	⋮	⋮	⋮	⋮
68	管理用イベントA	160.34	160.66	69	管理用イベントB	160.67
69	管理用イベントA	165.47	165.83	—	—	—
—	—	—	—	70	管理用イベントB	168.49
70	管理用イベントA	169.94	170.32	71	管理用イベントB	170.32
71	管理用イベントA	170.77	171.15	72	管理用イベントB	171.13
—	—	—	—	73	管理用イベントB	171.85
72	管理用イベントA	172.27	172.64	74	管理用イベントB	172.67

【図20】



【図21】



【書類名】 要約書

【要約】

【課題】 開始時点および終了時点が異なる同じ楽曲の音声データに対し、正しいタイミングで演奏データの同期再生を開始可能な記録装置、再生装置、記録方法、再生方法、およびプログラムを提供する。

【解決手段】 コントローラ部6は、音声データの再生に合わせて行われるピアノ31を用いた演奏のデータを、SMFに記録する。その際、コントローラ部6は、楽曲の開始時点および終了時点の近辺に対応する音声データの一部を用いて参照用の音声データを生成し、参照用の音声データをSMFに記録する。続いて、コントローラ部6は音声データの再生に合わせて、SMFに記録されたデータの再生を行う。その際、コントローラ部6は、音声データを用いて判定用の音声データを生成し、SMFに記録された参照用の音声データと判定用の音声データとの相関性に関する指標を算出し、その指標に基づいてSMFのデータの再生タイミングを調整する。

【選択図】 図1

出願人履歴情報

識別番号 [000004075]

1. 変更年月日 1990年 8月22日

[変更理由] 新規登録

住 所 静岡県浜松市中沢町10番1号

氏 名 ヤマハ株式会社